Universidad Carlos III de Madrid
www.uc3m.es

TESIS DOCTORAL

Essays on Economic Inequality

Autor:

Iacopo Morchio

Director/es:

Andrés Erosa Etchebehere

Tutor:

Andrés Erosa Etchebehere

DEPARTAMENTO DE ECONOMIA

Getafe, Julio 2015

Universidad Carlos III de Madrid
www.uc3m.es

(a entregar en la Oficina de Posgrado, una vez nombrado el Tribunal evaluador , para preparar el documento para la defensa de la tesis)

TESIS DOCTORAL

ESSAYS ON ECONOMIC INEQUALITY

Autor: *Iacopo Morchio*

Director/es: Andrés Erosa Etchebehere

Firma del Tribunal Calificador:

Firma

Presidente: (Nombre y apellidos)

Vocal: (Nombre y apellidos)

Secretario: (Nombre y apellidos)

Calificación:

Getafe, de de

Acknowledgements

I am very grateful to my advisor Andrés Erosa, and also to Javier Fernández-Blanco, Matthias Kredler and Josè-Victor Rios-Rull. Their advice, guidance and support have been crucial in the development of this thesis and in my development (past and future) as an economist.

This dissertation has benefited from the comments and suggestions of many other people, including Sevi Rodriguez Mora, Pedro Gomes and Ludo Visschers. I extend my sincere gratitude to them as well.

My life as a graduate student has been fantastic thanks to Silvia Martinelli, whose love and support during my PhD have been one of the most important parts of my life, and to my friends and colleagues Federico Curci, Mehdi Hamidi Sahneh, Lovleen Kushwah, Salvatore Lo Bello, Federico Masera, Alessandro Peri, Omar Rachedi and Marco Serena, whose company I enjoyed greatly both at the university and (especially) outside.

I also want to thank my great colleagues Lian Allub, Peter Eccles, Andrés Garcia-Suaza, Emre Ergemen, Pedro Sant'Anna, Victor Troster and Nora Wegner for all the discussions about economics and not. I really hope we will meet again!

Finally, I would like to thank my family for all they have done for me in the past 28 (almost 29) years of my life. Thanking them for this thesis would be very reducing as I owe them almost everything I am today.

Abstract

The last years have seen a surging interest in inequality in our society and in the world, with particular emphasis on economic inequality. Both long-run trends and the recent economic crisis have contributed to an increase in the gap between the rich and the poor, raising new questions on why this happened and on whether society should (or shouldn't) counteract these forces. The three main chapters of this dissertation aim at understanding the reasons behind inequality in the probability of being unemployed, inequality in test scores influenced by parental investment, and persistence of employment status across generations.

Chapter 1 studies the concentration of lifetime unemployment and its determinants. Using panel data from the US, I document three new stylized facts on unemployment. First, 10% of workers account for two-thirds of unemployment in prime-age. Second, young unemployment predicts prime-age unemployment. Third, differences in job-finding rates between the most unemployed and the rest increase over the life cycle, while differences in separation rates shrink. I show that a model of heterogeneity across workers and information frictions, in which agents learn workers' types from their labor market history, is quantitatively consistent with all these facts. I find information frictions to be responsible for the whole decrease in job-finding rates of the most unemployed workers over the life cycle. The concentration and persistence of prime-age unemployment are mainly explained by heterogeneity across workers, while information frictions have a negligible role.

Chapter 2 focuses on the relation between returns to parental investment and the way parents decide to spend their time. Time allocation choices of households and test scores of children exhibit regular patterns in the data. First, households spend more than twice as much time in child care when their child is under 6 years old than in later years, and reduce labor supply by almost 10 % in the same years. Second, college-educated parents spend more time with their children than noncollege ones. Third, cognitive test scores are intergenerationally correlated. In order to analyze the mechanisms behind these facts, I build a model of parental choices and both cognitive and noncognitive skills formation,

which embeds the technology of skills formation estimated by Cunha, Heckman and Schennach (2010). I find that the model, jointly with the properties of the technology, can account qualitatively and quantitatively for the bulk of such patterns. I also use the model to simulate the effect of applying the German scheme of child allowances to the US economy. Consistently with the empirical literature, such a policy is found to have little effect on the intergenerational persistence of cognitive skills.

Chapter 3, jointly developed with Salvatore Lo Bello, investigates the relation between the employment status of parents and the employment prospects of offspring. We study how parental links affect employment prospects, using monthly job histories from the BHPS. We motivate our empirical strategy by means of a stylized model of intergenerational transmission of networks. We find that having the father employed rather than unemployed increases the employment rate by about 8 p.p. and the monthly job finding probability by at least 50% (5-6 p.p). The effect is even larger when the father and the offspring work in the same occupational group. The empirical evidence suggests that such results are due to informational advantages rather than human capital transmission, direct hiring or common shocks.

Contents

List of Tables

List of Figures

Chapter 1

Information Frictions, Match Quality and Lifetime Unemployment

1.1 Introduction

Using data from the 1979 National Longitudinal Survey of Youth, I document three novel facts on lifetime unemployment. First, two-thirds of observed prime-age unemployment between 1985 and 2010 is accounted for by 10% of workers[1]. Such concentration is due to both lower job-finding rates and higher job-separation rates of the most unemployed workers. Second, time spent in unemployment when young is a powerful predictor of time spent in unemployment during prime-age. By means of regression analysis, I show that this is not due to observable heterogeneity such as education, occupation, or health. Third, I show that the 10% most unemployed workers and the rest start their careers with similar job-finding rates, and that the job-finding rate of the most unemployed declines over the years while the one of the rest of workers stays relatively constant. Instead, differences in monthly job-separation rates[2] shrink: they start as large as 4 percentage points at age 20 and descend to two percentage points at age 35.

Why are separation rates so heterogeneous and persistent over the life cycle? Why does the job-finding rate of the most frequently separated workers decline? And why do the same workers experience both low finding rates and higher separation rates? The fact that those with a low job-finding rate tend

[1] This is true even within education-sex subgroups. Bad luck alone cannot explain why unemployment is so concentrated: the standard search-and-matching framework is at odds with this fact, because it features too many transitions in and out of employment for the majority of workers.

[2] Conditionally on facing a separation, I find that the likelihood of experiencing a firing/layoff/temporary job ended/quit does not vary between lifetime unemployment groups. Thus, this is not because one group was more frequently fired than the other, for instance.

to have a high separation rate is crucial to account for heterogeneity in lifetime unemployment: since unemployment is a nonlinear function of both, theories that account only for one or the other cannot reproduce the concentration of unemployment observed in the data.

The challenge is to build a theory that is consistent with all the micro facts presented above. I propose a directed-search model which succeeds in this regard. In the model workers can be of two types, high and low. A worker's type is initially unobserved by all agents in the market, who are allowed to learn workers' types from labor market histories. This feature allows the model to be consistent with the fact that, while differences in job-finding rates increase over the career, differences in separation rates become smaller. In the model, this is because workers who face frequent separations when young progressively find fewer jobs, but keep sampling jobs until they find one in which they are productive.

Information is symmetric: at the start of a worker's career, no agent in the market (including the worker herself) knows her type. Search is directed in the sense that workers decide to search for a job delivering a certain lifetime value. Upon matching with a firm, workers draw match quality from type-specific distributions, which is constant for the whole duration of the match. Firms write complete contracts and matches are destroyed whenever their surplus is negative. Match quality is an experience good as in Jovanovic (1979); output of a match is unobserved until a shock is realized, upon which output becomes known to the firm-worker match. Then, the match is kept or destroyed, leaving the worker unemployed in the latter case. Past realizations of match quality for each worker are observed by the market, which updates the probability that the worker is of high type accordingly. The probability of being high-type formalizes a notion of "résumé" based on the worker's labor market history: low values of match quality will lead the market to believe that the worker is more likely to be of low type, while high values will have the opposite effect. Thus workers' types are slowly learned from labor market histories and workers with different résumés apply to different submarkets. To the best of my knowledge, this is the first model in which job-finding rates, job-separation rates and the speed of learning are all endogenously determined at the same time, because workers are allowed to choose the submarket in which they search, understanding that each lifetime values entails a different job-finding rate, and the outside option of the worker evolves dynamically with her history determining both her future desired lifetime value and her probability of separating from a job.

Heterogeneity in lifetime unemployment comes from three sources in the model. First, it can be the result of bad luck, because any given type might draw low match qualities, which will ultimately lead to separations. Second, it can be the result of information frictions, that is low-type workers wrongly infer from lucky draws that they are likely to be good types, leaving their current jobs in order to sample a better job, only to experience a bad draw later and get unemployed again.

I estimate the model using data from the NLSY/79. The model is very successful in reproducing the observed concentration and persistence of unemployment, as well as the patterns of job-finding rates, job-separation rates and wages over the life cycle. The model delivers concentration of unemployment because low-type workers have a higher probability of drawing low-quality matches than high-type workers, and have a lower expected productivity; thus such workers face a higher separation rate and a lower job-finding rate at every age. It delivers persistence because low-type workers tend to experience frequent separations both when young and when prime-age, and job-finding rates that decline with age as the market recognizes them as low-type workers. Information frictions are crucial to match the life-cycle patterns of job-finding and job-separation rates by unemployment groups. I argue that a model based on human capital, rather than information frictions, would be inconsistent with these patterns because it would have the counterfactual implication that differences in separation rates increase over the life-cycle.

I calibrate an array of competing models and find that neglecting heterogeneity across workers makes it impossible to match the concentration and persistence of unemployment observed in the data. While uncertainty in match quality draws helps in matching the life-cycle profile of job-separation rates and the concentration of unemployment, heterogeneity across workers is crucial to match the persistent differences in job-finding and job-separation rates across workers I document. Furthermore, uncertainty in match quality draws is important because it slows down learning: if there is no uncertainty and workers only differ in mean match quality, learning is too fast and it is impossible to match the progressive decrease in job-finding and job-separation rates by prime-age unemployment groups.

Information frictions play an important role in the first part of workers' lives. In a quantitative exercise, I shut down information frictions and show that they are responsible for the entire decline in monthly job-finding rates of the top 10% of prime-age unemployed (from 23 % at age 20 to 17 % at age 35). This is because 94% of the top 10% unemployed are low types: while their type is initially unknown, it is slowly revealed by their labor market histories. This translates in progressively lower job-finding rates for these workers. Information frictions also explain a portion of the decline in the separation rates of the most unemployed workers. However, the role of information frictions later in life is negligible: by age 30 types have effectively been learned for most of the population, so that most of the concentration and persistence of unemployment after this age are due to heterogeneity across workers.

This paper mainly contributes to two strands of the literature. First, it relates to the large empirical literature that investigates the composition of the unemployment pool and heterogeneity in job-finding rates; Clark et al. (1979) were the first to show that most of unemployment is accounted for by workers experiencing long spells of unemployment, rather than by workers going in and out of unemployment. In this paper, I make a different point and argue that most of the prime-age unemployment pool is composed by a relatively small group of workers continuously going out of employment, and staying unemployed for a long time, during all their lives. The literature on lifetime unemployment is relatively scarce, possibly due to the limited availability of long panel data. My results on the concentration of unemployment in the US are mirrored in the empirical work of Schmillen and Moller (2012), who use long time series from administrative data from Germany, and in Brooks (2005), who looks at workers in Canada in the years 1993-2001. Neither of these studies compares concentration to what is implied by standard models of unemployment. My approach at studying inequality in unemployment risk is similar to the one used in Michelacci, Pijoan-Mas and Ruffo 2011: using NLSY/79 data, the authors show that unemployment over the lifetime is more unequally distributed than what the standard search and matching framework implies. However, none of the studies above looks at the concentration in prime-age unemployment, nor documents young-prime-age persistence, nor decomposes concentration into job-finding and job-separation rates. A vast literature studies heterogeneity in job-finding rates and unemployment duration, both empirically (Addison and Portugal 1989) and theoretically (Lockwood 1991; Shimer 2008; Gonzalez and Shi 2010; Fernàndez-Blanco and Preugschat 2014; Wiczer 2014).

Second, I develop a model of unemployment and learning from job histories in which wages, job-finding rates, job-separation rates and the speed of learning are all jointly determined in equilibrium; this is also the first model to be estimated on (and to study) lifetime unemployment data. Other models of job search have proposed learning as a candidate explanation for the scars of unemployment (Michaud 2014) and duration dependence in job-finding rates (Gonzalez and Shi 2010). The model I develop shares a mechanism similar to Michaud (2014) regarding separations, but adds résumés, learning from labor market history and heterogeneity in the shape of match quality distributions across types: I find that all these ingredients are important to match heterogeneity in lifetime unemployment. My model's environment is similar to Menzio and Shi (2011), but in their model workers are heterogeneous in their ability to find jobs, and learn about their ability by finding jobs or not. Instead, in my model workers are heterogeneous in productivity and learn from their employment history, which maps into differences in job finding rates and job-separation rates. Similarly to Gonzalez and Shi (2010), my model also features a duration dependence relation because workers who have a higher probability of being high types tend to find jobs faster. Also, this paper shares with theirs having learning in a directed search framework, although Gonzalez and Shi model learning about the workers' ability to find jobs while I model learning workers' types (and thus, their distribution of productivity draws) from their labor market outcomes. My results on the speed of employer learning are similar to those of Lange (2007), who finds that employers learn relatively quickly and expectation errors on productivity decline by 50% in the first 3 years of employment. Other empirical work focuses on employer learning as a source of increase in wage heterogeneity over the career: see for instance Kahn and Lange (2013).

This paper also relates to the empirical literature that looks at the effect of unemployment on subsequent earnings and labor market outcomes. Since the pioneering study by Heckman et al. (1980), many papers have addressed whether unemployment leaves "scars" on subsequent wages and increases chances of future unemployment; see for instance Von Wachter, Manchester and Song (2009), Von Wachter and Bender (2006), Barnett and Michaud (2012). See also Couch and Placzek (2010) for a review of the studies on the effects of job displacement on earnings. Other recent studies (Kahn 2010; Oreopoulos et al. 2012) look at individuals who graduated from college during a recession, and find that this has negative, persistent effects on the earnings of otherwise identical workers. My model generates ex-post heterogeneity in labor market outcomes by allowing the market to separate workers using their job history, and in principle could be extended to allow for other additional channels discussed in the literature.

Finally my model is, in spirit, a life-cycle model of search and matching. Menzio, Telyukova and Visschers (2012) is the closest model to the one presented in this paper, having in common directed search and job-specific match quality. However, while they want to provide a life cycle theory of the transitions in and out of unemployment and employment over the life cycle, I want to understand the sources of heterogeneity in lifetime unemployment, and study the role played by information frictions in determining lifetime outcomes. I find that models without heterogeneity and learning (like Menzio, Telyukova and Visschers 2012 or Chèron, Hairault and Langot 2013), despite featuring potential sources of persistence such as human capital accumulation, cannot replicate the amounts of concentration and persistence of unemployment I document.

1.2 The Data

I use weekly job histories from NLSY/79 data to compute lifetime unemployment statistics. The NLSY is one of the best-known panel datasets available for the US, following a cohort of more than ten thousand individuals from 1979 onwards. Those who are being followed in the NLSY/79 ranged ages 14 to 22 in 1979; information has been gathered annually until 1994, and biennually since then.

I use only the cross-sectional representative sample of the NLSY, and exclude every worker who has less than 100 weeks of reported employment/unemployment from age 20 to 30, and 100 weeks from age 35 to 55; this gives me a sample of 5422 workers[3]. Further, I restrict attention to the relatively narrower sample of males who are only high-school educated at age 30[4]. This leaves us with a total of 1029 individuals followed for 30 years. However, results are robust to more inclusive definitions of the sample[5].

1.2.1 Prime-age unemployment is concentrated

I first document that prime-age unemployment is concentrated in relatively few individuals. I start by defining young-age unemployment as the fraction of the work history an individual spent in unemployment, over total weeks employed or unemployed[6], from age 20 to 30:

$$\bar{u}_i^y = \frac{\sum_{t=1}^{T_i^y} u_{i,t}^y}{T_i^y} \tag{1.1}$$

where $u_{i,t}^y$ is a variable taking value 1 in weeks in which individual i was unemployed, and 0 if individual i was employed, and T_i^y is the number of weeks that individual i was either employed or unemployed between ages 20 and 30. Similarly, I define prime-age unemployment as the fraction of work history spent in unemployment from age 35 to 55. Since I will show that there are important connections between young and prime-age unemployment, the five-years gap is necessary in order to

[3]This is to address measurement error issues when computing lifetime unemployment statistics. I study the extent of measurement error in Appendix 1.B.3.

[4]This means that I include in the sample only individuals who have completed no more and no less than high-school at age 30. I do this to have as homogeneous a sample as possible. High-school males are the biggest sex-education subgroup in the NLSY/79. Moreover, Menzio, Telyukova and Visschers (2012) show that, in terms of labor market outcomes, this subgroup is a good representation of the behavior of US labor market aggregates over the life cycle. In appendix 1.B.2 I show that findings are robust to other education-sex subgroups.

[5]For results on the whole sample, see Appendix 1.B.2.

[6]My definitions are similar to Schmillen and Möller (2012).

avoid that part of the correlations are not artificially due to the aftermath of a recession, or to long unemployment spells that connect between subsequent years.

As shown in table 1.1, there are large differences in unemployment outcomes across workers. The first finding is that prime-age unemployment is concentrated in relatively few workers. After ranking individuals by the fraction of time spent in unemployment, I compute the fraction of weeks spent in unemployment by the bottom 90% of the sample[7]:

$$\bar{u}^p_{u^p<q_{90}} = \frac{\sum_{i=1}^{N} 1(\bar{u}^p_i < q_{90}(u^p)) \sum_{t=1}^{T^p_i} u^p_{i,t}}{\sum_{i=1}^{N} 1(\bar{u}^p_i < q_{90}(u^p)) T^p_i} \tag{1.2}$$

where $1(u^p_i < q_{90}(u^p))$ is an indicator function taking value 1 if prime-age unemployment of individual i was below the 90th quantile of the prime-age unemployment distribution, and 0 otherwise, while T^p_i is the number of weeks in which individual i was either employed or unemployed during prime-age[8].

The 10% most unemployed individuals account for about 2/3 of prime-age unemployment observed in the data. Moreover, about half of these individuals have never been unemployed in the reference period. Notice that the fact that prime-age unemployment is concentrated in relatively few workers is a very different point from the one raised for instance by Clark et al. (1979), who show that most of the unemployment pool is accounted for by workers staying unemployed, rather than workers going in and out of unemployment.

I then proceed to compute monthly average job-finding/job-separation rates for workers in their primes[9]. I find that the concentration of unemployment is both due to a ($\simeq$ 3 times) lower finding rate[10]

[7]These measures are common in the literature on income inequality; see for instance Atkinson (1970). Their application to lifetime unemployment is relatively uncommon, with the exception of Schmillen and Möller (2012) and Brooks (2005).

[8]Clearly this is not the only way of computing this average. Another possibility is to compute instead

$$\tilde{u}^p_{u^p<q_{90}} = \frac{\sum_{i=1}^{N} \left[1(\bar{u}^p_i < q_{90}(u^p)) u^p_i \right]}{\sum_{i=1}^{N} 1(\bar{u}^p_i < q_{90}(u^p))} \tag{1.3}$$

that is, the average of each individual's prime-age unemployment. The two averages are different since T^p_i differs across individuals, because some are observed for more weeks than others; in particular, there can be a significant difference if $\mathbb{COV}(T^p, u^p) \neq 0$, for instance if those often unemployed tend to be more often out of the labor force. In fact, this is indeed the case (see Appendix 1.B.1). I find that there is relatively little difference between the two ways of computing the average, and that this does not matter for results on the concentration of unemployment, which is even larger (about 70% accounted for by top 10%) if using this second methodology (see Appendix 1.B.1).

[9]Since I will calibrate the model to monthly probabilities, I do not adjust for short-term unemployment as in Shimer (2012).

[10]I describe how I compute job-finding and separation rates in appendix 1.7.

	NLSY/79	Unif. Match	
		300 wks	500 wks
Avg. % time in unemployment	3.6	(target) 3.6	(target) 3.6
Avg. % time in U, excluding top 10%	1.5	2.5	3.1
Avg. % time in U, excluding top 20%	0.6	1.9	2.6
% never unemployed:	56	29	5

Table 1.1: Left column: averages computed on NLSY/79, individuals aged 35-55. Sample includes only high school educated, male individuals with more than 100 observations of weekly job histories in their prime-age, ending 2010. Right columns: averages computed by simulating sequences of 300 (column 2) and 500 (3) job-finding - job-separation events using flow equations of Mortensen-Pissarides model, calibrated to average job-finding and job-separation probabilities in NLSY/79 sample.

and a ($\simeq$ 9 times) higher separation rate for that top 10% (see table 1.2); this group of workers appears to have both longer unemployment duration and shorter employment duration. Since unemployment is a nonlinear function of both finding and separation rates, failure to account for both at the same time means not getting the distribution of unemployment right. Interestingly, the difference in separation rates accounts for a larger fraction of the heterogeneity in unemployment outcomes than the difference in finding rates.

	Top 10%	Rest of Sample	Ratio Top 10 / Rest
Avg. % time in unemployment	29	1.5	19.1
δ: Prob. of U $\rightarrow$ E (monthly%)	8	26	0.3
f: Prob. of E $\rightarrow$ U (monthly%)	3.5	0.4	8.75
Predicted % time in U of top 10%, δ alone:	12		
Predicted % time in U of top 10%, f alone:	5		
Avg. Log Wage (2000)	$\simeq$ - 40%		

Table 1.2: Summary statistics by parts of the prime-age unemployment distribution. Source: own calculations on NLSY/79. Male, high-school educated individuals aged 35-55. Predicted % time in U calculated using the formula $u = \delta/(\delta + f)$.

Similarly to what happens when discussing income inequality, measures of concentration might not be meaningful if they are not compared with what a standard framework would imply for the distribution of unemployment. If only one person out of 10000 was unemployed, the fact that unemployment is concentrated would not be very interesting. Moreover, it is important to stress that these numbers do not represent accurately differences in the "underlying" job-finding and job-separation rates for groups of workers. My estimates of job-finding and job-separation probabilities are likely to be biased estimates of the underlying probabilities, because by creating groups based on the amounts of unemployment experienced in prime-age I am selecting those individuals who experienced exceptionally high amounts of unemployment, who might be the most "unlucky" among a specific group. In order to understand the magnitude of these results, I compare the concentration of unemployment observed in the data to what a standard search and matching framework à-la Mortensen and Pissarides (1994) would imply. I produce simulations of 300 and 500 weeks of transitions because in my NLSY/79 sample I observe prime-age workers for about 700 weeks on average; 95% of workers are observed for more than 470 weeks, and less than 1 % of workers is observed for less than 250 weeks. This is for robustness: increasing the number of simulated weeks leads to worse performance of the standard model, so I construct at least one case that is favorable to it.

Simulations show that the standard model, calibrated to reproduce the job-finding and job-separation rates of the sample, has trouble replicating the observed concentration in prime-age unemployment: the standard search model features too many transitions in and out of unemployment for the majority of workers. This fact is important, because it suggests that heterogeneity across workers is likely to be crucial to make sense of labor market outcomes, and of the ins and outs of unemployment, during prime-age.

1.2.2 Unemployment is persistent over the life-cycle

I now document that young and prime-age unemployment are strongly correlated. Workers who were in the top 10% of the young-age distribution are five times more likely to be in the same part of the distribution when prime-age. In short, young and prime-age unemployment are connected and, among a wide range of observables available in the NLSY/79, young unemployment is the best predictor of prime-age unemployment. Noticeably, regression analysis (see table 1.7 in the Appendix) confirms that young unemployment is a very strong predictor of prime-age unemployment, and that this is not due to observables such as education, marital status or IQ.

Little additional information can be obtained by decomposing further the separation rate: using

All sample

	Rest	Top 10% when (35-55)
Rest	92.68	7.32
Top 10% when (20-30)	65.94	34.06

High-School Workers

	Rest	Top 10% when (35-55)
Rest	93.53	6.47
Top 10% when (20-30)	58.82	41.18

Table 1.3: Markov Transition Matrix, from distribution of young (ages 20-30) unemployment to prime-age (35-55) unemployment. Overall sample (top panel) and only high school males (bottom panel). Source: own calculations on NLSY/79.

the matched employer-employee dataset available along the NLSY/79, I show that workers who were in the top 10% most unemployed in prime-age had about twice the likelihood of separating from their employers for any reason than the rest of the sample, without one single particular reason being more important than others (see table 1.13 in Appendix).

Finally, notice that such persistence is not due to observable heterogeneity: one might think for instance that such differentials could be explained by differences across occupations (the choice of an "unlucky occupation" when young, as in Schmillen and Möller (2012)) or in health status (worse health means worse labor market outcomes). I perform several batteries of regressions (see Appendix 1.B.4) including ethnic origin, education, prior occupations and current occupations, ex-post health and IQ and find that none of these variables substantially reduces the amount of persistence I observe in the data. This result is particularly strong because current occupations and ex-post health are endogenous to prior labor market experience, and as such are bound to capture part of the persistence of unemployment. For instance, a worker that has been unemployed often when young will tipically work in more unstable occupations in prime-age, and this should capture part of the young-prime-age correlation I find. Similar considerations are valid for ex-post health.

1.2.3 Job-finding and job-separation over the life cycle

As a final piece of evidence, I compute job-finding and job-separation probabilities depending on age, from age 20 to age 35, by groups of prime-age unemployment. I want to show that those who have experienced large amounts of prime-age unemployment had different labor market outcomes during the first years of their career too. I compute marginal effects from linear regressions of job-finding rates and job-separation rates on a 4-th degree polynomial on age, controlling for year-specific fixed effects in order to clean the effect of recessions[11]. I can see that, at ages 20-30, the job-separation rate of the top 10% of prime-age unemployed is 4 percentage points higher than the job-separation rate of the rest of the sample (higher than the sample average), and this difference declines to 2 p.p at age 35. Instead, between the two groups there is only a 4 percentage points difference (about 1/7 of the sample average) in job-finding rates at age 20, but this difference becomes more pronounced as workers age, particularly because of the decline in the job-finding rate of the top 10% of prime-age unemployed.

Figure 1.1: job-finding (left panel) and job-separation (right panel) probabilities, by group of prime-age unemployed. Sample of male, high-school educated workers. Source: own calculations on NLSY/79. Shaded areas are 95% confidence bands.

This suggests that, in the eyes of potential employers, the two groups of workers were not substantially different at the beginning of their working careers, because they were hired with similar probabilities, but such differences became more pronounced later[12]; however, the high separation rates experienced by the top 10% of prime-age unemployed during their 20s suggest that such workers were recognized to be different *during* an employment relationship. That is, before an employment relationship had been established, young workers who came to experience substantially different careers looked similar;

[11] Results are substantially identical if I compute the averages using 5-years long age groups instead of restricting to a functional form. I choose the polynomial shape for presentation purposes; results under the age-group specification will be used to identify the model and will be reported in figure 1.4.

[12] I address why human capital-based explanations are insufficient to explain such patterns in section 1.6.

however, as they accumulated jobs and separations, workers experienced increasingly different job-finding rates, suggesting that information on them had slowly become available.

The wages of the top 10% unemployed progressively fall over the life cycle, relatively to those of the rest of the population (see figure 1.2), confirming that differences across workers become larger over workers' careers. This suggests that, after many separation events, such workers may sort into different jobs in order to avoid frequent future separations, or that they may fail to accumulate skills that lead to higher wages.

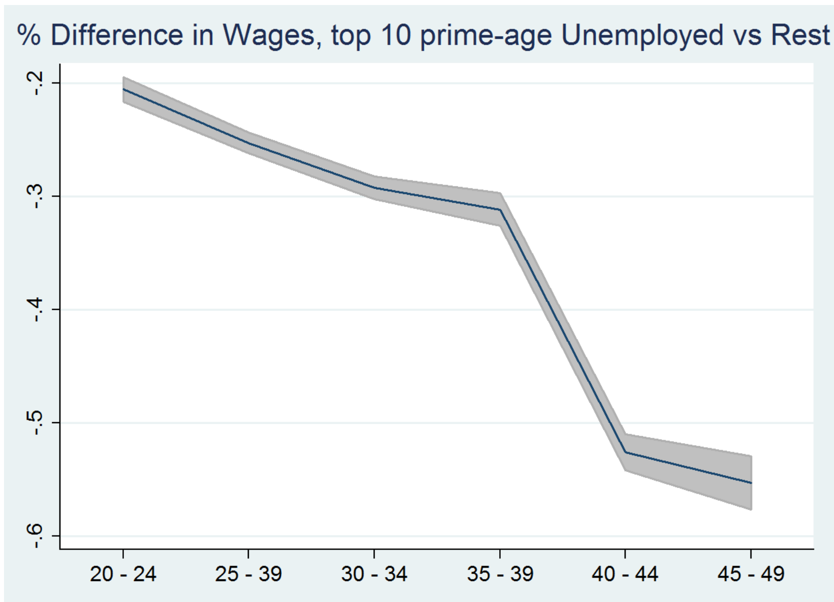

Figure 1.2: Log difference in Hourly Wage, between Top 10% Prime-Age Unemployment Group and Rest. Sample of male, high-school educated workers. Source: own calculations on NLSY/79. Shaded area is 95% confidence bands.

These facts motivate the need for a theory of unemployment that is capable of replicating the concentration of unemployment in relatively few workers and the persistence of unemployment over the life cycle; such concentration and persistence can have important consequences for the design of labor market policy. For instance, the concentration of unemployment suggests that relatively few people are likely to obtain the bulk of unemployment insurance, and will be the most affected by it. However, I argue that the relatively low ex-ante difference in job-finding rates and the large ex-post differences in both job-finding rates and wages suggest that important information frictions are at work in the first years of workers' careers, and that workers are being slowly sorted by employers over their careers. My model will feature this mechanism, which has important implications for understanding the concentration of unemployment, the connection of young and prime-age unemployment, and the effects of labor market policies.

1.3 Model

I now proceed to set up a model of heterogeneity in labor market outcomes, roughly based on Menzio and Shi (2011) and Gonzalez and Shi (2010); the ingredients of such model are inspired by the evidence presented in the previous section.

In order to obtain believable life-cycle profiles of separation rates, I add heterogeneity in match quality draws as in Menzio, Telyukova and Visschers (2012). Heterogeneity across workers, information frictions, and a notion of 'résumé' of the worker are added in order to capture the fact that a group of workers experiences higher separation rates at the start of the career, and that such separation rate diminishes later. This can be because such workers are being separated often (similarly to Gibbons and Katz 1991) and are learning that they have low productivity, thus their outside option decreases and reduces their probability of separating from a job. Moreover, such workers progressively find less jobs and earn lower wages: this can be rationalized by the fact that their résumé gets worse over their career, thus reducing their expected productivity in the eyes of potential employers. Moreover, heterogeneity across workers can rationalize the high levels of persistence of unemployment found in the data.

1.3.1 Environment

The economy is populated by a measure of firms $M > 1$ and a measure one of workers, who are either employed or unemployed. Every period, a fraction λ of workers die, and are replaced by newly born, unemployed workers. Each worker is born either of type H or L, High and Low respectively, unknown both to firms and workers; low types occur with probability l, high types with probability $1 - l$. All agents are risk neutral and discount the future at rate $1/(1 + r)$.

Let p be the probability of a worker being high-type. There exists a continuum of submarkets indexed by $\{v, p\}$, the expected lifetime value v earned by the worker in that submarket and the prior p of workers applying to that submarket[13]. Matches are endogenously destroyed when the surplus of the match is negative. Some matches end randomly with probability δ.

[13]This is to make less assumptions on the distributions of match quality that follow. In principle, the model can be rewritten to feature submarkets indexed only by $\{v\}$, provided that further assumptions on the match quality distributions are made so that workers with different p will apply to different submarkets. In the current version, I allow workers with different p to apply to the same lifetime value, but in equilibrium for every value of $p \in [0, 1]$ only one submarket $\{v(p), p\}$ will be active.

1.3.2 Search and Matching

Firms can post vacancies in any submarket at cost κ. Search is directed, in the sense that workers with prior $\bar{p}$ can choose in which submarket $\{v, \bar{p}\}$ to search. Thus, each submarket has tightness $\theta(v, p)$, the ratio of vacancies to searching workers. The number of matches in each submarket is determined by the matching function $m = g(\theta)$, such that the job-finding probability is $f(\theta) = m/u$, which satisfies $f' > 0$, $f'' < 0$, $g(0) = 0$ and $\lim_{\theta \to \infty} = 1$, and the job-filling probability is $q(\theta) = m/v = f(\theta)/\theta$. Unemployed workers can search for a job while employed workers cannot. When unemployed, workers get benefit b.

1.3.3 Information and Learning

Denote by $H(x)$ and $L(x)$ the cumulative distribution functions of match quality, for high and low types respectively, with support $X \subseteq [0, \bar{x}]$, such that $H(x)$ strictly first order stochastically dominates $L(x)$; that is, $H(x) < L(x)\ \forall x < \bar{x}$. Once a match with a worker has been established, a match-specific quality shock is drawn from the workers' type distribution. Match quality is constant over the whole duration of the match[14].

At the beginning of a firm-worker match, output of the match is unobserved. Match quality is an experience good as in Jovanovic (1979): after a match with a worker with belief p has been established, in each period the match produces expected payoff[15] $\mathbb{E}(x \mid p)$, until the worker/firm pair gets to observe output or a random separation occurs. With probability $1-\pi$, output is not observed. With probability π, output of the match is observed.

Upon observing output, agents gain information on a worker's type: higher output realizations signal that the worker is more likely to be of high-type, while lower output realizations signal the opposite.

1.3.4 Contracts

I assume that employment contracts are complete, in the sense that they specify a wage w paid by the firm to the worker and a probability of separation d at every point in time, as a function of the promised expected lifetime utility, the history of the worker and that of the firm-worker match[16]. As Menzio and

[14] Menzio, Telyukova and Visschers (2012) find that, in a similar model, the probability that a match changes quality during an employment relation is around 1%, thus making the constant match assumption a reasonable simplifying approximation.

[15] From the perspective of writing the surplus of a match, this is equivalent to assuming that both parties get zero value until productivity is observed and then get the sum of flow utilities of all previous periods.

[16] This assumption reflects the idea that matches can be kept as long as they are profitable to both parties, so that the relation between labor market histories and learning is not strongly dependent on the contract environment, but on

Shi (2011) prove, since contracts are complete and utility is transferable, it is optimal for firms to offer contracts that are bilaterally efficient, so that they maximize the sum of the firm's lifetime profits and the worker's lifetime utility. Thus, the firm finds optimal to offer a probability of separation d delivering bilateral efficiency (matches are kept as long as their lifetime value is higher than the outside option of the worker) and a wage w such that the expected lifetime utility v is delivered to a worker in a $\{v, p\}$ match. However, there are many different sequences of w that deliver the same lifetime utility to the worker; in this paper, it is not important to solve this indeterminacy as the object of discussion are the patterns of job-finding and job-separation rates[17].

The intuition is that jobs will be endogenously destroyed whenever the value of the match is lower than the outside opportunity of the worker: thus, low-type workers will face more frequent separations if they typically draw match qualities that are below the value of unemployment. Thus, π and the properties of the match quality distribution $H(x)$ and $L(x)$ measure the *informational content of job duration.* If $\pi = 0$, match quality is never observed and job duration is not informative of the worker's type.

Since the history of past match qualities drawn by a worker is observable by potential employers, it follows that p is a sufficient statistic for the whole history of match qualities a worker experienced, and can be considered the worker's 'résumé'.

Timing is as follows:

1. Workers die w.p. λ, replaced by unemployed workers with belief $1 - l$.

2. W.p. π, firms and workers observe output of a match.

3. Workers revise beliefs: $p' = p$ if match quality is still unobserved and no shocks occur, $p' = P(x, p)$ if match quality is observed today.

4. Production occurs, wages are paid.

5. Unemployed workers search for a job. They choose to search in submarket $\{v', p'\}$.

6. Workers match w.p. $f(\theta(v', p))$. Separations (exogenous and endogenous) occur.

the features of match quality distributions. The idea that types are being learned over workers' careers does not hinge necessarily on the particular contract space assumed here, but its quantitative implications might change. For instance, it is possible to think of an environment in which match quality is the firm's private information and workers learn from whether they are kept or fired by the firm.

[17] When thinking about how to map wages in the model to the data, it can be useful to rewrite the equilibrium definition including wages as equilibrium objects. I will discuss this issue in more detail in the Equilibrium subsection.

7. Newly matched workers draw match quality from $H(x)$ or $L(x)$ depending on their type.

Bayes' rule implies that beliefs of employed workers who observe the realization of match quality evolve according to

$$p' = P(x,p) = \frac{p\,h(x)}{p\,h(x) + (1-p)\,l(x)} \tag{1.4}$$

where $h(x)$ and $l(x)$ are the density functions of match quality draws for high and low types respectively.

Heterogeneity in job duration across workers and the update of p are closely related: when a worker is observed and the match continues, it must mean that match quality was high enough to support the match, an event that is more likely for high types because $H(x)$ first order stochastically dominates $L(x)$. Viceversa, when a worker's productivity is observed and the match is destroyed, it must mean that match quality was not high enough, an event that is more likely for low types. Clearly, this depends on how high the flow utility of unemployment is with respect to the tipical match quality draws of low types. However, depending on the features of the match quality distributions, matches might be destroyed also because workers prefer to have the opportunity to sample a new job with a higher productivity.

1.3.5 Bellman Equations

The value function of an unemployed worker with prior p can be written as

$$U(p) = b + \beta\left[\max_{v}\left[f(v,p)(v - U(p))\right] + U(p)\right] \tag{1.5}$$

where $\beta = \frac{1-\lambda}{1+r}$.

The joint firm-worker value of a match which output is known can be written as

$$S(x,p) = x + \beta\left\{\max_{d\in[\delta,1]}\left[(1-d)\,S(x,p) + d\,U(p)\right]\right\} \tag{1.6}$$

Thus, the joint value of a match in which output is unknown and a worker of prior p is working can be written as

$$\begin{aligned} S_u(p) = \quad & (1-\pi)\Big[\mathbb{E}(x \mid p) + \beta\Big((1-\delta)\, S_u(p) + \delta\, U(p)\Big)\Big] + \\ & \pi\Big(p\int S(x,p')\, dH(x) + (1-p)\int S(x,p')\, dL(x)\Big) \end{aligned} \tag{1.7}$$

where I denote by $p' = P(x,p)$ the next period's prior depending on the realization of match quality, suppressing the dependence on p and x for readability.

The value of a firm posting a vacancy in submarket $\{v,p\}$ is

$$V(v,p) = -\kappa + q(\theta(v,p))\beta(S_u(p) - v) \tag{1.8}$$

and the tightness function must satisfy

$$\kappa \geq q(\theta(v,p))\beta(S_u(v,p) - v) \quad \forall v,p \tag{1.9}$$

which makes θ consistent with the firm's optimal vacancy creation; 1.9 holds with equality if $\theta > 0$. Basically, condition 1.9 implies that if $\theta = 0$, such tightness is consistent with the firm's optimal choice only if the benefit from creating a vacancy is smaller than the cost.

1.3.6 Equilibrium

Definition 1. *a Bayesian Markov Perfect BRE (Block Recursive Equilibrium) for this economy consists of a value function for the unemployed worker $U(p)$, a policy function for the unemployed worker $v'(p)$, value functions for the joint value of a match $S_u(p)$ and $S(x,p)$, a separation policy $d(x,p)$, a tightness function $\theta(v,p)$ and a law of motion for beliefs $p'(x,p)$ such that*

1. $U(p)$, $v'(p)$, $S(x,p)$, $S_u(p)$, $d(x,p)$, $\theta(v,p)$ are independent of the aggregate state ψ
2. $\theta(v,p)$ satisfies 1.9 $\forall$ v,p and $\theta(v,p) \geq 0$ with complementary slackness.
3. $U(p)$ and $v'(p)$ satisfy 1.5 $\forall$ p
4. $S(x,p)$ and $d(x,p)$ satisfy 1.6 $\forall$ w,x
5. $S_u(x,p)$ satisfies 1.7 $\forall$ w,x

6. $p'(x,p)$ satisfies 1.4

The BRE is much easier to solve than a Recursive Equilibrium, because value functions and policy functions of agents depend only on the individual states $\{p,\, x\}$ and not on aggregate states. Aggregate statistics can be computed, after solving the individual problem, from the aggregation of individual choices. Moreover, computing transitions out of the steady state is easy because all policy functions and laws of motion are independent of the aggregate state.

1.3.7 Characterization of Equilibrium

Lemma 1. *Given $p \in [0,1]$ and $x \in [0,\bar{x}]$, $S(x,p)$ is increasing in x. Also, $d(x,p)$ is a step function that takes value 1 if $x < U(p)(1-\beta)$ and δ otherwise.*

The intuition is simply that match quality must be high enough to be higher than the value of unemployment. Clearly, distributions of match quality that have more probability mass on values that are lower than b will lead to more frequent separations.

Lemma 2. *In the BRE of the economy, the unique solution to equilibrium condition 1.9 is*

$$\theta(v,p) = \begin{cases} q^{-1}(k/(\beta\,(S_u(p)-v) & \textit{if } \;\beta\,(S_u(p)-v) \geq k \\ 0 & \textit{otherwise} \end{cases} \tag{1.10}$$

Since the function $S_u(p)$ is continuous in p for $p \in [0,\,1]$, the market tightness function θ is continuous in p. Furthermore, since $V(v,p)$ is a decreasing function of v, $\theta(v,p)$ is a decreasing function of v. The intuition is that, as firms have to pay higher wages in order to deliver the promised level of lifetime utility, their expected profits are lower, so that a higher job filling probability is required to pay for the cost of creating a vacancy, thus implying a lower tightness in that submarket.

Remark: $P(x,p)$ is continuous in x and p $\;\forall x < \bar{x}$. Moreover, $\partial P/\partial p > 0$.

Continuity of posterior belief P and the fact that it is increasing in p stems trivially from the functional form. However, first order stochastic dominance is not sufficient to ensure that beliefs are monotone in x, which would require further restrictions on the shape of the distributions.

Lemma 3. *The optimal choice of a worker can be written as the choice of a tightness $\theta^*(p)$, $\forall\, p \in [0,1]$, and is unique given p.*

Proof. Notice that, by using the fact that $f(\theta)/q(\theta) = \theta$, equation 1.5 can be rewritten as:

$$U(p) = b + \max_{\theta}\left\{ f(\theta)\beta(S_u(p) - U(p)) - \kappa\theta \right\} + \beta U(p) \tag{1.11}$$

by using equation 1.9 for submarkets with a positive tightness.

Given p this is a well-defined, concave problem in θ because by the assumptions on the matching function $f(\theta)$ is an increasing, concave and twice-differentiable function. Thus, the derivative of the value function with respect to θ is

$$\frac{\partial U}{\partial \theta} = f'(\theta)\beta(SU(p) - U(p)) - \kappa$$

which admits a zero for $\theta > 0$. □

Lemma 4. *Under the assumption that $H(x) < L(x) \quad \forall x \in [0, \bar{x})$, U is strictly increasing.*

The complete proof can be found in appendix 1.C. The intuition is that higher résumés p will translate into higher expected productivity, thus to a higher expected job-finding rate for every promised lifetime utility, thus to a higher value of unemployment. This also means that workers will trade off higher job-finding rates with lower promised lifetime utility when choosing the submarket in which to search.

1.4 Quantitative Model and Identification

The model is identified by estimating parameters in order to replicate features of job-finding rates, job-separation rates and wage patterns observed in the NLSY/79. The idea behind identification is that the concentration and persistence of unemployment, and the differences in job-separation and job-finding rates by parts of the prime-age unemployment distribution, are informative of the amount of low-type workers present in the economy and of the differences between the match quality distributions of types, while the life-cycle profile of wages are informative of differences both in the match quality distributions and in skill multipliers. This strategy is partly inspired by Menzio and Shi (2011) and Menzio, Telyukova and Visschers (2012), who use the life-cycle patterns of job-finding rates, job-separation rates and employment-to-employment transitions in order to identify the parameters of the match quality distribution and the probability of observing productivity during a match. My model works similarly during a match's duration, so I apply the same strategy but I distinguish between the job-finding/separation rates experienced by the top 10% prime-age unemployed and the rest of the population.

To see why the match quality distribution affects separation rates, consider the separation policy

$d(x,p)$. Given a résumé p, the probability that a firm destroys a match upon descovering match quality is $H(U(p)(1-\beta))$ for high types and $L(U(p)(1-\beta))$ for low types, which means that the way in which the probability mass is distributed over match qualities determines separation rates for different types, which are also heterogeneous in their labor market histories.

Turning to how the match quality distribution affects job-finding rates, consider equation 1.9, which states that in equilibrium the tightness of submarket $\{v,p\}$ depends on the expected surplus of the firm for a worker with prior p. In turn, the expected surplus depends on $\mathbb{E}(x \mid p)$ and on $S(x,p)$, that is on current expected productivity and on future productivity if the match will not be destroyed. Basically, job-finding rates depend on expected match quality given the prior, and on the expected match quality for the part of the distribution above the separation cutoff.

Summing everything up, a distribution featuring high mass on low values of match quality, but a long right tail, will deliver high separation rates and high job-finding rates. On the other hand, a distribution featuring high mass on low values of match quality and a short right tail will deliver high separation rates and low job-finding rates. Finally, a concentrated distribution, such that uncertainty about match quality is low, will deliver low separation rates.

I now explain why the concentration and persistence of unemployment is informative on the amount of low-type workers and the match quality distributions. Consider a case in which workers have the same starting résumé p (the population prior), and the match quality of low types has more probability mass on low realizations than the match quality distribution of high types. This means that young, low-type workers who are starting their careers will typically experience a larger-than-average amount of separations during their youth. As information on their type accumulates, these workers will slowly sort into lower-wage jobs, but will still experience higher separation rates because of the worse match quality distribution, and will experience lower job-finding rates because their expected productivity will be lower. The mechanism does not necessarily apply only to low types: high types who have been unlucky, and drew many low-quality matches, will experience frequent separations and will be considered "low types" with a high probability, thus experiencing lower job-finding rates when older.

Young-age separation rates depend on how fast output is observed (parameter π), while the speed of learning depends on how far apart the two distributions of match quality are: if types draw match qualities from very similar distributions, learning will be slow, job finding rates will be similar and the concentration of unemployment will be low too. If the two types draw from very different distributions, learning will be fast and unemployment will be concentrated in few workers. The scale and shape of the match distributions will thus influence the life-cycle profile of job-finding rates, job-separation rates and wages. Notice that it is possible to obtain concentration of unemployment even with only one type of workers, just by changing features of the match quality distribution. However, this would be inconsistent with the fact that unemployment is persistent over the life cycle, and with the life-cycle patterns of job-finding rates and job-separation rates by unemployment groups.

Persistence of unemployment depends on how far apart the two distributions of match quality are, how risky they are and how large the measure of low-types is. If low-types have always high risk of being unemployed (that is, of drawing low match quality values) while high-types are almost never unemployed, persistence will be high and will be determined almost uniquely by the measure of low-types. To see why, suppose that low-types are 10% of workers: in that case, persistence as measured by the probability of ending up being in the top 10% of prime-age unemployment, given that one has been in the top 10% of young unemployment, will be 100%. However, there will be no additional persistence at the top 20% because the rest of the population is never unemployed both when young and in prime-age.

Instead, if the two distributions of match quality are very close, persistence of unemployment will depend also on how fast learning is, and on the role of luck in determining unemployment for both types. In all cases, however, the persistence of unemployment over the life-cycle can be used to pin down the measure of low-types present in the economy.

1.4.1 Calibration

I now proceed to simulate the lives of a large sample of workers in order to compute lifetime statistics, and calibrate the model to replicate as closely as possible the observed patterns of wages and transition rates in the NLSY79. Estimation is performed by applying the Simulated Method of Moments: I minimize the loss function

$$L(\omega) = m(\omega)' \, W \, m(\omega) \tag{1.12}$$

where ω is the vector of parameters of the model, $m(\omega)$ is a column vector of the differences between the model-generated moments and the data moments, and W is a weighting matrix[18].

[18] Computation of variance-covariance matrix of moments and of standard errors is not trivial, because one moment restriction comes from the estimates of Hall and Milgrom (2007) and its covariance with the remaining moments cannot be computed. At the moment, W is set in such a way that moments are scaled to their data average, that is I minimize the sum of the square differences $\frac{m(\omega)}{\hat{m}}$, where $\hat{m}$ are the data moments: in this way, I minimize the sum of relative distances from data averages.

I set the model period to be one month. I assume workers are born at age 20, the starting age of my data, and choose the death probability λ in order to match an average working life of 40 years. I choose the interest rate r as to give a compounded annual interest rate of 4%.

In line with many other models of directed search (Shimer (2005); Mortensen and Nagypal (2007); Menzio and Shi (2011); Menzio, Telyukova and Visschers (2012)), I restrict the matching probability to be of the form $f(\theta) = \min\{\theta^{0.5}, 1\}$.

The flow value of unemployment b is considered as including both the value of leisure and unemployment benefits, and is chosen as to match a ratio between b and average wages[19] of 0.71, in line with the estimates of Hall and Milgrom (2007).

The two match quality distributions H and L are assumed to be Weibull distributions[20] with scale parameters σ_H, σ_L and shape parameters ϕ_H, ϕ_L. Shape and scale of match quality distributions, the probability π of observing a worker's output, the random separation probability δ and the measure of low-type workers l, are calibrated to match the observed patterns of job-finding rates, job-separation rates over the life cycle by rest of population and top 10% unemployed, and the observed concentration and persistence between young and prime-age unemployment of top 10% and of top 20%, as in the tables presented in section 1.2. Notice however that the model is unit-free, so one of the scales has to be set exogenously. I normalize $\sigma_H = 1$.

The vacancy creation cost κ is calibrated as to match the job-finding rate of bottom 90% of the prime-age unemployment distribution at ages 20-25.

The calibration table reports only singleton targets: patterns of job-finding/separation rates and wages are vectors and are shown later in graphs for readability. Overall, the estimation algorithm fits 8 parameters with 30 restrictions.

1.5 Results

1.5.1 Calibration results

Despite being calibrated with over-identifying restrictions, the model does a very good job in replicating the main features of the data. As can be seen in table 1.4, the model is quite capable of delivering realistic

[19]This requires to compute wages implied by the choices of workers in terms of lifetime utility v. For calibration purposes, I assume that wages are determined by the firm as the constant piece-rate of productivity (that is, a constant fraction of productivity) that delivers the promised expected lifetime utility v (provided the match continues).

[20]The Weibull distribution is a common choice in this regard. See for instance Menzio, Telyukova and Visschers (2012).

Parameter	Symbol	Value	Target	Data	Model
Interest rate	r	0.0034	Yearly discount rate = 0.96		
Prob. of Death	λ	0.0021	Avg. career length = 40 yrs		
Scale param, high type	σ_H	1.0000	JS rate profiles,		
Scale param, low type	σ_L	0.4335	JF rate profiles,		
Shape param, high type	ϕ_H	15.4297	% prime-age U acc. by top 10	0.60	0.46
Shape param, low type	ϕ_L	1.3051	% prime-age U acc. by top 20	0.83	0.68
Prob. observing output	π	0.0558			
Flow value of Unemp.	b	0.5384	Ratio b/avg. wage	0.70	0.65
Vacancy Creation Cost	κ	2.9225	JF rate, bottom 90, age 20-25	0.29	0.31
Random Sep. Rate	δ	0.0042	JS rate, bottom 90, age 40-50	0.007	0.007
Measure of Low Types	l	0.2722	Persistence of U, top 10	0.41	0.39
			Persistence of U, top 20	0.45	0.45

Table 1.4: Baseline calibration results. Targets calculated on NLSY/79.

amounts of concentration and persistence of unemployment. The model fits very well the persistence as measured by the Markov transition matrix between being unemployed when young and when prime-age: the probability of being in the top 10% of the unemployment distribution when prime-age, after having been in the top 10% of the unemployment distribution when young, is 0.39 in the model and 0.41 in the data. At the top 20%, the same statistic is 0.45 in the model and 0.45 in the data.

The model falls a bit short in replicating the observed concentration of the distribution of prime-age unemployment: the top 10% accounts for 46% of prime-age unemployment in the model and 60% the data, while the top 20% accounts for 68% of prime-age unemployment in the model and a bit more than 80% in the data. However, this is a drastic improvement with respect to the standard Mortensen-Pissarides model, which only predicts one-third of observed concentration at the top 10%.

The match quality distributions of high- and low-type workers are substantially different: at the calibrated values, the match quality distribution of low types has more mass close to zero, and a long right tail, while the match quality distribution of high types is narrower and more concentrated on higher match qualities (figure 1.3).

The calibrated value of the probability of a firm observing the worker's output $\pi = 0.0558$ implies that the average duration of a "bad match" is about 18 months.

The calibrated measure of low-type workers in the economy is around 27%, a relatively large number.

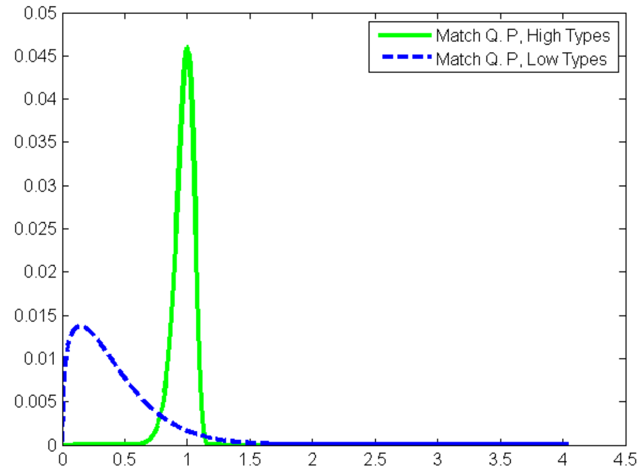

Figure 1.3: Match quality distribution of high types (red) and low types (blue), under baseline calibration

As I will show in the discussion section, this number has important implications for the composition of the unemployment pool and for the concentration and persistence of unemployment over the life cycle.

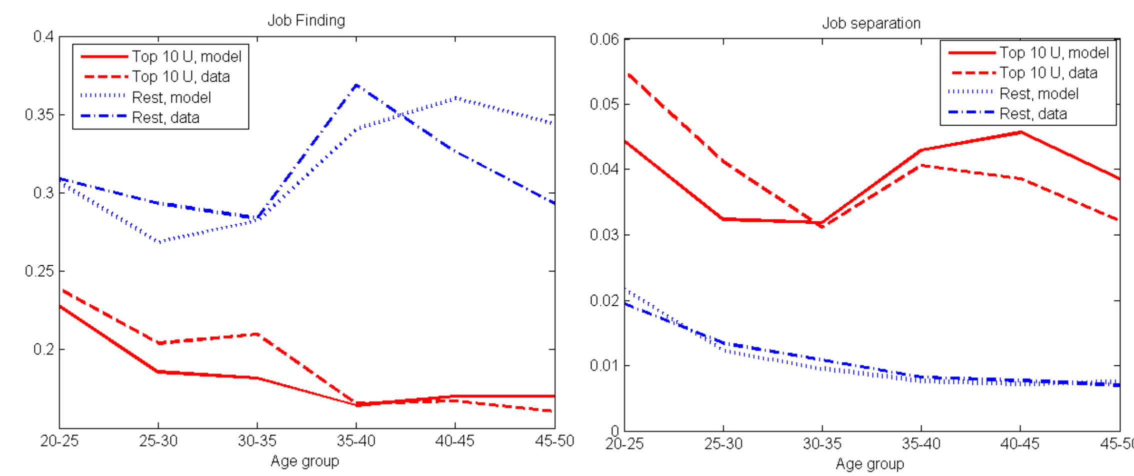

Figure 1.4: Model versus data; job-finding rates (left) and job-separation rates (right) of top 10% prime-age unemployed (red) vs rest (blue). Results under baseline calibration.

Figure 1.4 shows that the job-finding rate of the top 10% of prime-age unemployed declines over the life cycle as in the data, while the job-finding rate of the rest of workers rises during prime-age. The model is very successful in fitting the patterns of job-separation and job-finding rates by prime-age unemployment groups, both for the most unemployed and the rest.

Figure 1.5 plots the probability that a worker is of high-type depending on her age, by low and high types and by part of the prime-age unemployment distribution. The figure depicts what I term “learning over the life cycle”: as separations and continuations occur, the market slowly learns who are high-type and who are low-type workers. The patterns of job-finding and job-separation rates are a consequence of this mechanism.

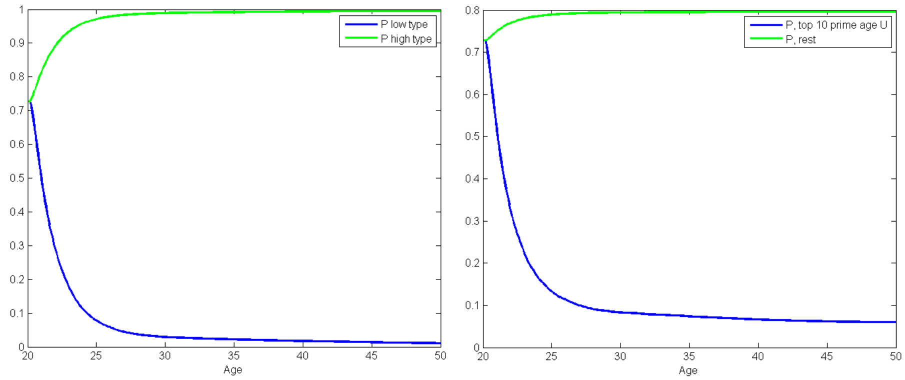

Figure 1.5: Probability of being a high type: by high/low type (left) and by top 10% of prime-age unemployment (right). Model results under baseline calibration.

Let us look first at the job-finding rate: as the market learns who are low- and who are high-type workers, the gap in job-finding rates between workers widens. This can be seen by comparing job-finding rates and job-separation rates of high and low-type workers in the model (figure 1.6). The result follows from this mechanism, and from the fact that more than 70% of the unemployment pool is made of low types (figure 1.10 in Appendix). Thus, the job-finding rate of the top 10% unemployed is essentially the job-finding rate of the most unlucky of low types: the model predicts that 94% of the top 10% unemployed in prime-age are low-type workers.

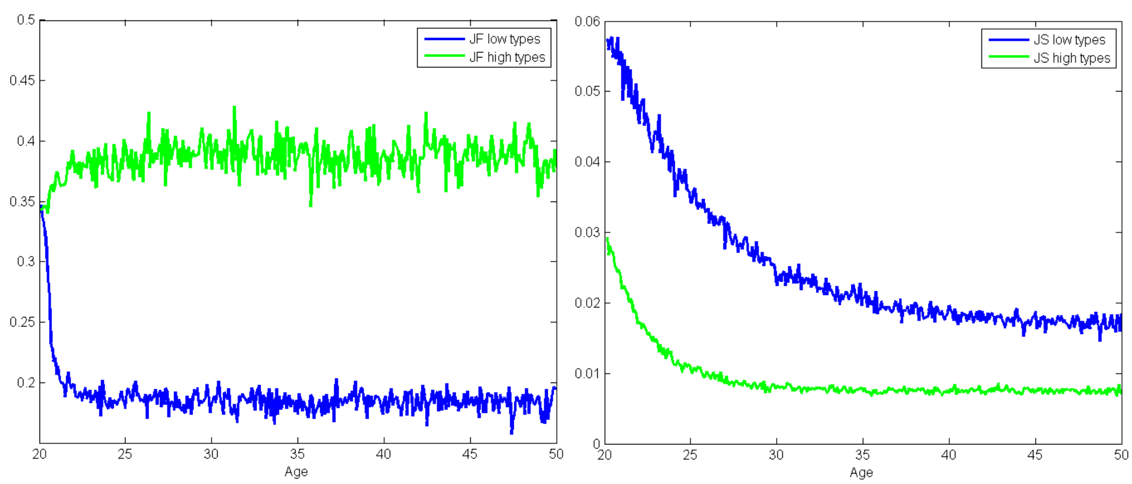

Figure 1.6: job-finding rates (left) and job-separation rates (right) of low types (blue) vs high types (green). Model results under baseline calibration.

Job-separation rates are substantially higher for the top 10% unemployed, both when young and when prime-age; the model replicates quite well both the descent in the job-finding rate of the most unemployed workers and the increase in the job-finding rate of the rest over the life-cycle. If anything, the model undershoots slightly the separation rate of the most unemployed workers until age 35, while it overshoots it a bit for the same workers later on.

Similarly to job-finding rates, the job-separation rate of low-type workers is affected by learning over the life cycle. At ages 20-30, the job-separation rate of low-type workers declines because these workers typically draw low values of match quality and face frequent separations. However, both workers and the market learn from these separations, so that the workers' outside option declines and these workers become progressively less likely to separate from a job. Moreover, a luck effect exists: sooner or later, every worker can find a job in which she is productive and stay there. The subsequent rise in separation rates observed for the top 10% of prime-age unemployed is due to selection bias: this empirical strategy is selecting the most unemployed individuals, who tend to be the most unlucky of low-type workers.

1.5.2 Counterfactual Simulations

I now simulate what would happen in alternative scenarios, removing model features one by one to study their relative importance in fitting the data. Results are summarized in table 1.5. All models have been recalibrated on the same loss function of the baseline model.

First, I calibrate a version of the model featuring uncertainty about match quality, on-the-job human capital accumulation and stochastic human capital depreciation when unemployed (column 1): such a model can get close to the data in terms of concentration of prime-age unemployment (at the top 10%, 44% against 60% in the data) thanks to uncertainty in match quality draws. Moreover, this feature allows to replicate some of the life-cycle profile of separation rates: the top 10% prime-age unemployed start with a job-separation rate of 2.80 at age 20-25, against 5.51 in the data, and the separation rate at age 40 for these workers is almost matched (4.01 against 3.86). However, such a model fails completely in delivering sufficiently large persistence of unemployment (at the top 10%, 12% against 41%). This is because human capital accumulation and depreciation introduce "reshuffling" in the skill level of workers: instead of having fixed differences in productivity, every worker can now become unskilled if he stays unemployed long enough, or skilled if he manages to get a sufficiently high level of match quality. Moreover, model 1 cannot replicate the patterns of differences in separation rates at young ages by unemployment groups, because it does not feature enough heterogeneity in match quality draws: the bottom 90% of workers have about the same separation rate of the most unemployed at ages 20-30, differently from the 3.6% difference existing in the data. Finally, such model cannot replicate the fact

	(1) 1 type $\mathbb{VAR}(x) > 0$ HC	(2) 2 types $\mathbb{VAR}_l = 0$ $\mathbb{VAR}_h = 0$	(3) 2 types $\mathbb{E}(x\|l) =$ $\mathbb{E}(x\|h)$	Baseline	Data
Separation rates, top 10					
age 20-25	2.80	1.22	2.84	4.45	5.51
age 25-30	2.21	1.17	2.09	3.24	4.12
age 40-45	4.01	2.21	4.21	4.56	3.86
Separation rates, rest					
age 20-25	2.63	1.13	2.61	2.16	1.95
age 25-30	1.68	1.11	1.54	1.23	1.34
age 40-45	0.66	1.00	0.62	0.72	0.77
Finding rates, top 10					
age 20-25	24.06	26.00	24.14	22.79	23.88
age 25-30	23.61	18.19	23.33	18.57	20.41
age 40-45	18.72	14.65	18.77	17.02	16.73
Finding rates, rest					
age 20-25	24.01	32.38	24.17	30.66	30.90
age 25-30	24.15	30.66	23.98	26.85	29.33
age 40-45	28.86	36.85	29.02	36.04	32.62
Concentration - Prime-age					
U acc. by top 10	38.11	29.29	37.70	45.71	60.00
U acc. by top 20	59.91	46.58	65.96	68.21	83.00
Persistence					
Top 10 young - Top 10 primes	12.27	23.73	14.62	38.79	41.00
Top 20 young - Top 20 primes	25.19	26.99	24.26	44.82	45.00

Table 1.5: Baseline calibration results vs counterfactuals. Column 1 is a model with no unobserved heterogeneity, no uncertainty in match quality and fixed observable skills. Column 2 adds uncertainty in match quality and accumulation/depreciation of skills. Column 3 is a model with heterogeneity in average productivities, but no match quality uncertainty. Column 4 has heterogeneity in match quality uncertainty, but not in average productivities. All numbers are percentage points. All models have been recalibrated on the same loss function.

that the job-finding rate of the most unemployed falls at the beginning of their career.

I now calibrate a model with no human capital, heterogeneity in mean productivity across types, but no uncertainty about match quality (column 2): that is, I estimate a model forcing the distribution of match quality to be degenerate. Such model cannot predict the higher separation rate at age 20-25 for one group of workers, because as soon as workers draw their first job, they immediately learn that they are low types and give up looking for a better job forever. In synthesis, learning is too fast and bad luck plays little role: when there is no uncertainty about match quality, a worker will learn her type with certainty at the first observation of output. Moreover, since the distribution of match quality is degenerate, there is no other mechanism that delivers heterogeneity in separation rates. Such model also predicts a too sudden and too large decrease in job-finding rates, which start substantially higher than the data at age 20. This is another consequence of excessively fast learning, and of the struggle of the model in delivering concentration of unemployment by having to rely solely on heterogeneity in job-finding rates. Finally, this model can improve over the model without heterogeneity (model 1) in delivering persistence of unemployment (24% against 41% in the data at the top 10%), but fails dramatically in delivering concentration, underperforming even the model without heterogeneity (model 1) in this regard; this is because there is no uncertainty in match quality, and thus no heterogeneity in separation rates.

In the last experiment, I calibrate a model with heterogeneity in the variance of the match quality distribution, but no differences in mean productivities[21] (column 3). Such model performs very similarly to the model without heterogeneity, delivering too little heterogeneity in job-finding rates and job separation rates of young workers, while performing better for old workers.

These quantitative exercises confirm that all ingredients are important for explaining the patterns observed in the data. Heterogeneity in the mean of match quality draws is important for explaining differences in job-finding rates. Heterogeneity in the variance of match quality is important for explaining heterogeneity in job-separation rates, for obtaining concentration of unemployment and for slowing down learning at the start of the career: slower learning translates into a more realistic descent of job-finding rates and job-separation rates for the most unemployed workers. Notice that even a model with homogeneous types, but uncertainty in match quality draws (model 1), is capable of delivering concentration of unemployment: this is because such concentration can be obtained if there is sufficient heterogeneity in separation rates across workers, which can be the consequence of bad luck in match quality draws.

1.5.3 Decomposing Learning over the Life Cycle

In this section I keep the baseline calibration but shut down information frictions by making types known right from the start, to understand their importance in explaining the data. Results are presented in table 1.6

[21] In practice, this is done by letting the shape parameters ϕ_i of the Weibull distributions be estimated freely by the algorithm, while σ_l solves the nonlinear equation

$$\mu_h = \mu_l$$

where $\mu_i = \sigma_i \Gamma(1 + \frac{1}{\phi_i})$ is the mean of the Weibull distribution and $\Gamma(x)$ is the Gamma function.

	Types known	Baseline	% Difference
Separation rates, top 10			
age 20-25	3.90	4.45	-12.16
age 25-30	2.96	3.24	-8.38
age 30-35	3.11	3.19	-2.40
age 40-45	4.58	4.56	0.37
Separation rates, rest			
age 20-25	2.14	2.16	-1.13
age 25-30	1.19	1.23	-2.93
Finding rates, top 10			
age 20-25	18.67	22.79	-18.08
age 25-30	18.46	18.57	-0.60
age 40-45	16.88	17.02	-0.81
Finding rates, rest			
age 20-25	18.67	22.79	-18.08
age 25-30	18.46	18.57	-0.60
Concentration - Prime-age			
U acc. by top 10	45.36	45.71	-0.76
U acc. by top 20	68.11	68.21	-0.15
Persistence			
Top 10 young - Top 10 primes	39.17	38.79	0.97
Top 20 young - Top 20 primes	44.96	44.82	0.30

Table 1.6: Baseline results vs counterfactuals. Column 1: types are known from the beginning. Column 2: baseline model. Column 3: percentage difference between column 1 and 2. All numbers are percentage points.

First, I shut down entirely learning over the career, by making types already known at the beginning. This reduces the separation rate of the most unemployed at age 20-30 by about half a percentage point (one-eigth), because these workers are already aware that they are low types, so their outside option is already low and they do not separate from jobs after lucky draws. The persistence and concentration of unemployment are barely affected, as these are mainly due to heterogeneity across workers and bad luck in drawing match quality values. Finally, information frictions account for the whole decline in job-finding rates from age 20 to 40 for the most unemployed workers: if types were already known, firms would already anticipate their low average productivity and these workers would find jobs with lower probability right from the start.

1.5.4 Duration dependence

The model is also capable of reproducing a duration dependence relation in job-finding rates (figure 1.7), similar to the one documented by Hornstein (2012) and Wiczer (2014). The relation arises because of a composition mechanism similar to Gonzalez and Shi (2010): workers with higher market prior find jobs first, followed by workers with lower market priors. I plan to expand this section in the future by decomposing the duration dependence relation in effects of learning vs observable skills.

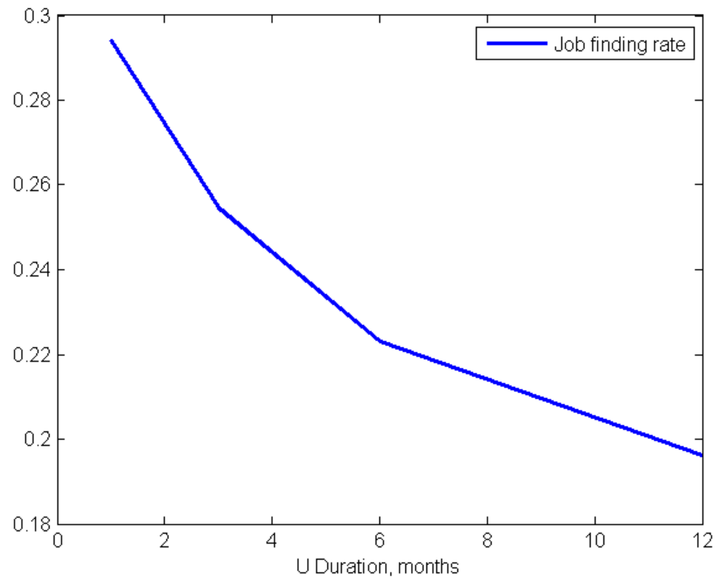

Figure 1.7: Model-generated data: duration dependence relation in job-finding rates, at 1, 3, 6 and 12 months of unemployment duration.

1.6 Discussion

1.6.1 Heterogeneity or Human Capital?

I have shown that a theory of information frictions and heterogeneity is capable of explaining at the same time the patterns of job-finding rates, job-separation rates and (part of) the patterns of wages by unemployment groups over the life cycle. An alternative explanation might be that workers who are often unemployed tend to lose, or fail to accumulate, human capital because they lack on-the-job training and face human capital depreciation (as in Ljunqvist and Sargent (1998)). To the extent that human capital is observable, if workers started with some level of human capital, depreciation would lead the most unemployed workers to experience lower job-finding rates, possibly explaining one of the facts. However, even if lower human capital yielded higher separation rates, depreciation would imply that heterogeneity in job-separation rates *rises* along the career, because the most unemployed would lose human capital and possibly face higher separation rates, while the rest of workers would experience fewer separations. As a result, we should observe a divergence in separation rates by unemployment groups, and not a convergence such as the one I document.

Column 1 of table 1.5 partially tests for these implications by calibrating a version of the model featuring no ex-ante heterogeneity across workers, match quality draws and human capital accumulation/depreciation as the only source of persistence in unemployment. Such model delivers almost no young-prime-age persistence of unemployment, and patterns of job-finding rates and job-separation rates by prime-age unemployment groups that are completely inconsistent with the data.

1.7 Conclusions

Using NLSY/79 data, I show that unemployment during prime-age is concentrated in relatively few workers, who experience both long spells of unemployment and frequent separations from their jobs. Moreover, unemployment is persistent in the sense that those who were often unemployed when young tend to be often unemployed during their primes. I build a model that delivers both high concentration of unemployment during prime-age and persistence of unemployment over the life-cycle, and that is consistent with the patterns of job-finding rates and job-separation rates by prime-age unemployment groups. The model delivers such result by a combination of incomplete information and heterogeneity across workers. I find that information frictions are important for explaining workers' labor market outcomes at the beginning of their career; in particular, a model without information frictions delivers a too high job-finding rate gap between different workers at the start of their work life and a higher separation rate than the one observed in the data for most young workers. Finally, I find that unobserved heterogeneity, rather than differences in observed skills, is responsible for the bulk of my results.

Appendix

1.A Construction of job-finding and job-separation probabilities

Following Clark and Summers (1979) and Wiczer (2014), I consider workers who exit the labor force as if they were not in the at-risk population; for each group of workers N_j, which can be the whole sample ($N_j = N$), or the top 10% of the unemployment distribution and its complement, I use the formulas

$$F_j = \frac{\sum\limits_{i \in N_j} \sum\limits_{t=1}^{U_i^p} f_{i,t}}{\sum\limits_{i \in N_j} U_i^p} \tag{1.13}$$

$$S_j = \frac{\sum\limits_{i \in N_j} \sum\limits_{t=1}^{E_i^p} s_{i,t}}{\sum\limits_{i \in N_j} E_i^p} \tag{1.14}$$

where $f_{i,t}$ is a variable defined only in weeks spent in unemployment, which were followed by weeks spent in either unemployment or employment, and takes value 1 if the following week the worker was employed, and 0 otherwise; $s_{i,t}$ is defined only in weeks spent in employment, followed by weeks spent in either employment or unemployment, and takes value 1 if the following week the worker was unemployed, and 0 otherwise; U_i^p is the number of weeks worker i was unemployed during prime-age; and E_i^p is the number of weeks worker i was employed during prime-age.

Dep. Variable: % U when 35-55

	(1) Only HS, Males	(2) All, with controls	(3) All, No Young U
% U when young (20-30)	0.3635***	0.239***	
	(0.020)	(0.02)	
N	1029	3127	3127
R^2	0.186	0.218	0.127

p-values in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.7: Source: own calculations on NLSY/79. Regression of % of prime-age unemployment on % of young unemployment: only for high-school educated males (1), for all workers + controls (2), for all workers with controls only and no young unemployment (3). Controls include sex, education, ethnic group, age in 2010, marital status, AFQT test score quartile.

1.B Supplementary Data Analysis

1.B.1 Impact of Labor Force Participation

In this subsection I investigate whether the labor force participation margin is relevant for the results I present on the concentration of unemployment. One possibility is that the most unemployed individuals get discouraged about their possibilities of finding jobs; thus, they might tend to permanently drop out of the labor force more frequently than the rest of the sample. To study whether this is the case, I compute the average participation rate of individuals by unemployment groups. It is easy to see that the top 10% of prime-age unemployed tends to participate less often to the labor force. The two groups follow a substantially parallel trend until age 40, after which the top 10% do tend to drop out of the labor force more frequently. However, when we look at individuals who did not participate for a full year, this difference reduces dramatically, suggesting that although the top 10% tends to spend more time out of the labor force, this does not mean that they always drop out completely.

It is unlikely that changes in sample composition are driving most results on the concentration of unemployment; however, I address this concern by studying how much the participation margin matters for computing lifetime statistics and the concentration of unemployment. As explained in section 1.2, another possible way to compute the average by unemployment groups is

$$\tilde{u^p}_{u^p<q_{90}} = \frac{\sum_{i=1}^{N}\left[1(\bar{u}_i^p < q_{90}(u^p))u_i^p\right]}{N} \tag{1.15}$$

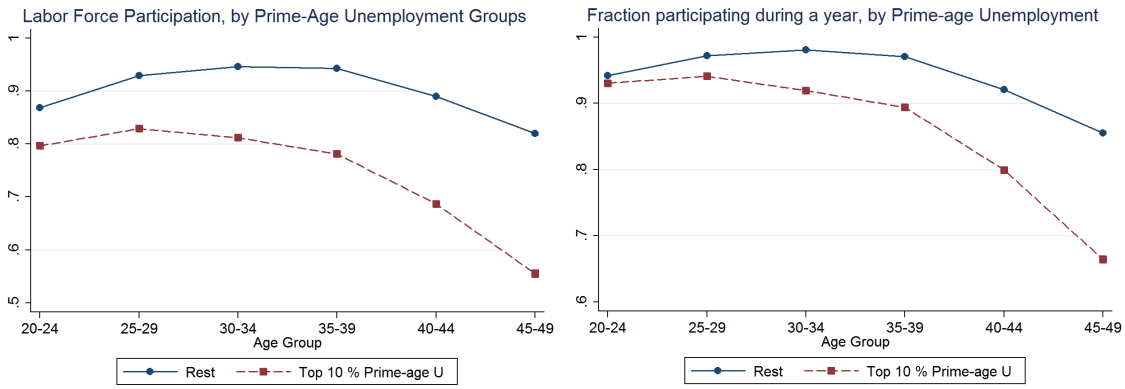

Figure 1.8: Labor Force Participation rate; by prime-age unemployment groups and by 5-years age groups. **Left**: fraction of weeks spent either in employment or unemployment by group. **Right**: fraction of individuals in group who worked at least a week during a year.

so that every individual has the same weight in the computation of the average, regardless of the number of periods he has been employed or unemployed. I will refer to this as the *equally-weighted* average, and to the average presented in the paper as the *participation-weighted* average.

In principle, it is not clear which of the two averages should be used. Since the top 10% of prime-age unemployed tends to be out of the labor force more often, these individuals have a lower weight in the participation-weighted average than in the equally-weighted average. Thus, the latter represents the concentration of unemployment if we were to observe the top 10% in the labor force as often as the rest of the sample. With respect to this logic, the participation-weighted average is likely to bias downward my estimates of aggregate prime-age unemployment, and of the concentration of unemployment. As shown in table 1.8, the equally-weighted formula indeed imply substantially identical averages when excluding the most unemployed, but a higher average of overall prime-age unemployment. Thus, this implies a higher concentration of unemployment, compared to the results with participation-weighted averages; the performance of the standard model is even worse in delivering concentration of unemployment when using equally-weighted averages for comparison.

1.B.2 Sample selection

In this subsection I show that sample selection plays little role in computing both the concentration and persistence of unemployment. I compare the statistics computed in section 1.2 with the same statistics[22] computed using the whole sample of workers by education and gender.

[22]In this case, the average time in unemployment during prime-age is the *participation-adjusted* average.

	NLSY/79	Unif. Match
Avg. % time in unemployment	4.6	(target) 4.6
Avg. % time in U, excluding top 10%	1.6	3.1
Avg. % time in U, excluding top 20%	0.6	1.9

Table 1.8: Left column: **equally-weighted** averages computed on NLSY/79, individuals aged 35-55. Sample includes only high school educated, male individuals with more than 100 observations of weekly job histories in their prime-age, ending 2010. Right column: averages computed by simulating sequences of job-finding - job-separation events using flow equations of Mortensen-Pissarides model, calibrated to average job-finding and job-separation probabilities in NLSY/79 sample.

	Only HS	Whole Sample
Prime-age unemployment Avg. % time in U	3.6	3.4
Avg. % time in U, exclud. top 10%	1.6	1.4
Avg. % time in U, exclud. top 20%	0.6	0.6
Persistence Prob. top 10 prime-age given top 10 young	41.18	34.06
Prob. top 10 prime-age given Rest young	6.47	7.32
Prob. top 20 prime-age given top 20 young	44.88	43.07
Prob. top 20 prime-age given Rest young	13.71	14.23

Table 1.9: **Participation-adjusted** averages computed on NLSY/79, individuals aged 35-55. Left column: only high school educated, male individuals with more than 100 observations of weekly job histories in their prime-age, ending 2010. Right column: whole cross-sectional sample of NLSY/79, satisfying the same restriction on weekly job histories.

I also investigate whether concentration and persistence of unemployment, as well as differentials in job-finding rates and job-separation rates I document, vary significantly across education subgroups. I keep only males and divide the NLSY/79 into high-school dropouts, high-school educated and some-college and above (those who took some college courses but did not complete college, and college-educated). Results are summarized in table 1.10. For all subgroups, all facts stand. Unemployment is more concentrated than what the Mortensen-Pissarides standard model implies; high young unemployment predicts high prime-age unemployment; and inequality is more due to heterogeneity in job-separation rates than job-finding rates.

	Dropouts	High-School	$\geq$ Some College
% U Accounted for by top 10	47	59	70
Predicted by MP	23	24	44
% U Accounted for by top 20	69	83	92
Predicted by MP	38	39	70
Persistence: prob. of top 10 prime-age			
from top 10 young	26	41	22
from rest young	7	6	7
Avg. % time in unemployment			
top 10% prime-age	53	29	15
Rest	4.2	1.5	0.5
δ: Prob. of U $\rightarrow$ E (monthly%)			
top 10% prime-age	4	8	11
Rest	18	26	35
f: Prob. of E $\rightarrow$ U (monthly%)			
top 10% prime-age	5.0	3.5	2.0
Rest	0.8	0.4	0.2
Predicted % time in U of top 10%, δ alone:	21	12	5
Predicted % time in U of top 10%, f alone:	16	5	2

Table 1.10: Summary statistics by parts of the prime-age (35-55) unemployment distribution and by education subgroups. Source: own calculations on NLSY/79. Predicted % time in U calculated using the formula $u = \delta/(\delta + f)$.

1.B.3 Measurement Error

When computing lifetime unemployment statistics, it is crucial to have enough observations per individual. If an individual had been observed only for few weeks, and was always unemployed, taking the average over those weeks would incorrectly attribute a lifetime unemployment of 100% to that individual. To address the extent of measurement error, I compute the concentration of unemployment in top 10 %, top 20 %, and the persistence of unemployment for different values of the lower bound of weeks of reported employment/unemployment, both when 20-30 and when 35-55. Results are reported only for the high-school subsample. Although measures of persistence and unemployment tend to fall, because the most unemployed also tend to stay out of the labor force more often, results on concentration are substantially unchanged, and persistence remains high: in the worst-case scenario in which the sample is required to have at least 500 weeks of reported employment/unemployment both when young and when prime-age (totaling 1000 weeks over 1560 maximum weeks available), the top 10% most unemployed when young still have between 4 and 5 times the likelihood of being the most unemployed when prime-age, and the most prime-age unemployed still account for two-thirds of unemployment.

	100 weeks (baseline)	300 weeks	500 weeks
Prime-age U	3.6	3.3	2.6
Prime-age U, without top 10%	1.5	1.3	0.9
Prime-age U, without top 20%	0.6	0.5	0.3
Persistence (top 10 - top 10)	41%	35%	29%
(rest - top 10)	6%	7%	8%
Persistence (top 20 - top 20)	45%	41%	35%
(rest - top 20)	14%	14%	16%
N. Individuals	1029	918	633

Table 1.11: Accounting for possible measurement error: concentration and persistence of unemployment according to alternative definitions of the sample. High school males with at least 100 weeks (column 1), 300 weeks (2), 500 weeks (3) of reported employment/unemployment. Source: own calculations on NLSY/79.

1.B.4 The role of Occupations and Health

One might think that differences across occupations are behind the strong young-prime-age correlations found in the data. For instance, the choice of a “bad occupation” when young might lead a worker to experience high unemployment both when young and in the future. I show that occupations explain relatively little of the observed young-old persistence by augmenting previous regressions with occupational controls (table 1.12). I use the CENSUS 1970 classification at the major category level, and I control both for the most prevalent occupation between 1979 and 2001 (that is, the occupation in which the individual worked the most during those years) and for the occupation in 1990. Sample size diminishes because occupation codes are not always available for workers in the NLSY/79; however, the strong predictive power of young unemployment remains substantially unchanged. Occupation in 1990 appears to be the most important correlate variable, diminishing the young-prime-age persistence of unemployment by 0.06. However, one must consider that occupational choice in 1990 is not independent on past labor market history, and its relevance is likely to be upward-biased because of reverse causality.

I finally consider whether the deterioration of health correlates with prime-age unemployment. I use ex-post health (in 2006) as a control as to construct a worst-case scenario: since health in 2006 can be the result of past unemployment, it will partly correlate with young unemployment and prime-age unemployment, thus in principle lowering the estimate of the impact of young unemployment. I find that, although health correlates with prime-age unemployment, it has a negligible influence on the predictive power of young unemployment.

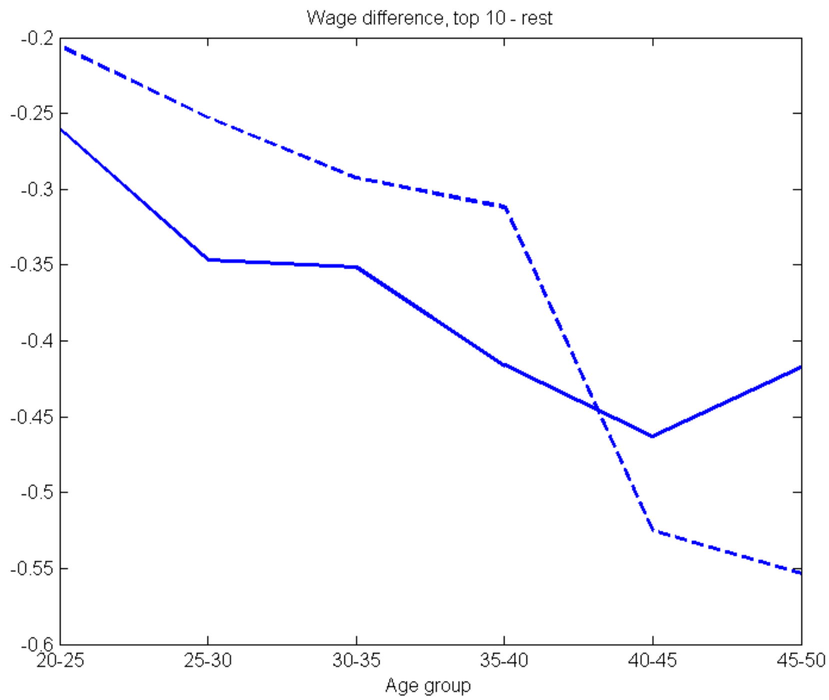

Figure 1.9: Difference in wages between top 10 % prime-age unemployed and rest; data (dashed) versus model (continuous) under baseline calibration.

	(1) Base	(2) Educ	(3) Educ+Occ	(4) Educ+Occ 2	(5) Ed+Occ 2+Hlth
% U when Young (20-30)	0.299***	0.256***	0.256***	0.254***	0.247***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Age	-0.000262	-0.000233	-0.000144	-0.000159	-0.000195
Female	0.00128	0.00211	0.000769	0.000910	-0.000363
Ethnic Group					
Black	0.0204*	0.0170*	0.0160*	0.0165*	0.0145*
Hispanic	-0.000845	0.00827	0.00744	0.00760	0.00623
Marital Status					
Married		-0.0340***	-0.0334***	-0.0332***	-0.0319***
Separated		-0.0139	-0.0130	-0.0132	-0.0141
Divorced		-0.0119*	-0.0111*	-0.0114*	-0.0119*
Widowed		-0.00665	-0.00637	-0.00631	-0.00508
Education, age 30					
Some College		0.00770	0.00477	0.00505	0.00429
High School		0.00929*	0.00578	0.00564	0.00289
Dropout		0.0295***	0.0268***	0.0265***	0.0230**
AFQT Quartile					
2		-0.00499	-0.00562	-0.00518	-0.00323
3		-0.0106*	-0.0108*	-0.00999	-0.00688
4		-0.0128*	-0.0122*	-0.0113	-0.00896
Health 2000					
Very Good					-0.00167
Good					0.00367
Fair					0.0196**
Poor					0.0685***
Constant	0.0403	0.0440	0.0369	0.0371	0.0373
Standard Controls	X	X	X	X	X
Education, AFQT and MaStat		X	X	X	X
Prevalent Occ. (1 digit)			X	X	X
Occupation in 1990 (1d)				X	X
Health Status (2000)					X
N	3896	3896	3896	3896	3896
R^2	0.151	0.179	0.183	0.185	0.195

Standard errors in parentheses
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 1.12: Source: own calculations on NLSY/79. Complementary regressions of % of prime-age unemployment on % of young unemployment for all workers. Sample restricted to individuals for which all controls are available for all models. Controls always include sex, ethnic group and age in 2010. (2) adds AFQT test score quartile, education and marital status, (3) adds prevalent occupation during working life dummies, (4) adds occupation in 1990 dummies, (5) adds health status in 2000 dummies. Omitted categories: male, white, never married, college-educated, 1st quartile AFQT, Technical/professional occupations, Excellent Health. Occupation coefficients and standard errors (with the exception of young unemployment) are not reported for reading convenience.

	Rest	Top 10% (35-55)	Ratio Top 10 / Rest
When young (20-30):			
Fired	0.06	0.14	2.33
Involuntary	0.43	0.78	1.8
Quit to Look	0.4	0.6	1.5
In Primes (35-55):			
Fired	0.02	0.1	5
Involuntary	0.09	0.4	4.44
Quit to Look	0.02	0.09	4.5

Table 1.13: Weekly probability of job termination, by reason and group of prime-age unemployment. Third column gives ratio of probability between top 10 and rest. Source: own calculations on matched employer-employee data of NLSY/79. 'Involuntary' category merges layoffs, establishment closures and temporary jobs ended.

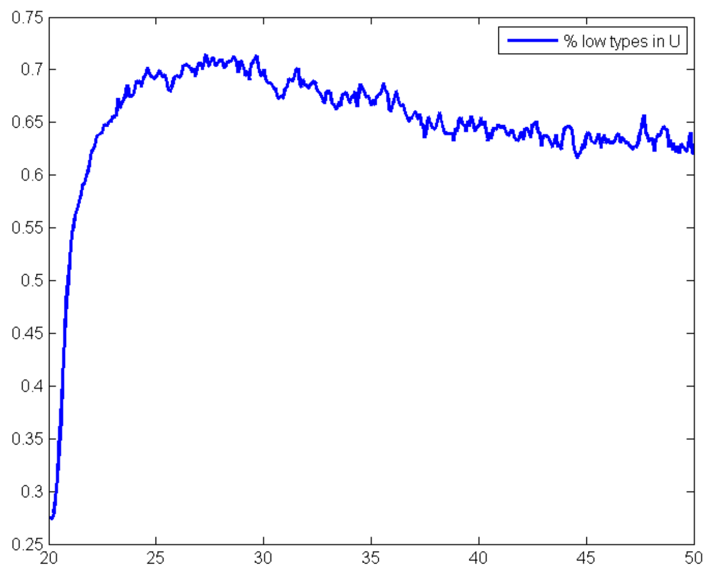

Figure 1.10: Share of workers who are low types, by age; under baseline calibration.

1.C Proofs not included in Paper

Lemma 1: Given $p \in [0, 1]$ and $x \in [0, \bar{x}]$, $S(x, p)$ is increasing in x. Also, $d(x, p)$ is a step function that takes value 1 if $x < U(p)(1 - \beta)$ and δ otherwise.

Proof: Given a match quality value x and a résumé p, the problem of a firm-worker pair can be rewritten as:

$$S(x,p) = \max_{d \in [\delta,1]} \left[\frac{x + \beta\, d\, U(p)}{1 - \beta\,(1-d)} \right] \tag{1.16}$$

which is clearly increasing in x. As both x and p are given, so is $U(p)$. By taking first derivative it can be shown that the function on the right hand side of the expression is increasing in d if and only if

$$x < U(p)(1-\beta) \tag{1.17}$$

Denote by d^* the optimal choice for d. If condition 1.17 is verified, maximization implies that $d^* = 1$. If instead $\frac{x}{1-\beta} > U(p)$, the function is decreasing in d and thus $d^* = \delta$ which is the lowest possible probability of separation. If $x = U(p)(1-\beta)$, the firm-worker match is indifferent between any value of $d \in [\delta, 1]$. $\square$

Lemma 5: Under the assumption that $H(x) < L(x) \;\; \forall\, x \in [0, \bar{x})$, U is strictly increasing.

Proof. Let $TU(p)$ denote the right hand side of equation 1.11 so that U is a fixed point of the mapping T. It must be proved that the operator T maps from the set of continuous, bounded, increasing functions (denoted by C_1) into the subset of itself containing continuous, bounded and strictly increasing functions C_1^s. Thus, it must be proved that $TU(p_2) > TU(p_1) \;\; \forall\, p_2, p_1 \in [0,1]$ such that $p_2 > p_1$. Since T is a contraction mapping, it will follow that U is strictly increasing. Denote by $\theta * (p)$ the optimal tightness choice of a worker with prior p. Hence,

$$
\begin{aligned}
& U(p_2) - U(p_1) \\
= \;& \max_{\theta}\Big\{ f(\theta)\beta(S_u(p_2) - U(p_2)) - \kappa\theta \Big\} + \beta U(p_2) \\
& - \max_{\theta}\Big\{ f(\theta)\beta(S_u(p_1) - U(p_1)) - \kappa\theta \Big\} - \beta U(p_1) \\
\geq \;& f(\theta^*(p_1))\beta\big[SU(p_2) - SU(p_1)\big] + (1 - f(\theta^*(p_1))\beta\big(U(p_2) - U(p_1)\big) \\
\geq \;& f(\theta^*(p_1))\beta\Big[SU(p_2) - SU(p_1)\Big] \\
= \;& B\Big\{(1-\pi)\big[\mathbb{E}(x|p_2) - \mathbb{E}(x|p_1)\big] + \beta(1-\pi)\delta\big[U(p_2) - U(p_1)\big] \\
& + \pi\Big[\Big(p_2\int S(x,p_2')dH(x) + (1-p_2)\int S(x,p_2')dL(x)\Big) \\
& - \pi\Big[\Big(p_1\int S(x,p_1')dH(x) + (1-p_1)\int S(x,p_1')dL(x)\Big)\Big]\Big\} \\
\geq \;& B\Big\{(1-\pi)\big[\mathbb{E}(x|p_2) - \mathbb{E}(x|p_1)\big] \\
& + \pi\Big[(p_2 - p_1)\int S(x,p_1')dH(x) + (1-p_2)\int S(x,p_2')dL(x) \\
& + (1-p_2)\int S(x,p_1')dL(x) - (1-p_2)\int S(x,p_1')dL(x) - (1-p_1)S(x,p_1')dL(x)\Big]\Big\} \\
= \;& B\Big\{(1-\pi)\big[\mathbb{E}(x|p_2) - \mathbb{E}(x|p_1)\big] \\
& + \pi\Big[(p_2 - p_1)\int S(x,p_1')dH(x) - (p_2 - p_1)\int S(x,p_1')dL(x) \\
& + (1-p_2)\big[\int S(x,p_2')dL(x) - \int S(x,p_1')dL(x)\Big]\Big\} \\
> \;& 0
\end{aligned}
$$

The first equality is obtained by substituting equation 1.11. The second inequality is because maximization implies that a worker's optimal choice $\theta * (p_2)$ will always deliver a higher or equal value than if he chose $\theta * (p_1)$ instead. The fact that U is increasing delivers the next inequality. The fourth equality is obtained by simply substituting the solution to equation 1.7, where $B = \frac{f(\theta^*(p_1))}{1-\beta(1-\pi)(1-\delta)}$ The fifth inequality stems again from the fact that U is increasing, so that also $S(x,p)$ is increasing in p, and from the fact that posterior beliefs p_i' are increasing in p, so that $\int S(x,p_2')dH(x) \geq \int S(x,p_1')dL(x)$; notice that at this step I add and subtract the same quantity $(1-p_2)\int S(x,p_1')dL(x)$. The sixth equality is obtained by regrouping terms; notice that at this point the first line is strictly positive due to First Order Stochastic Dominance, the second line is positive for the same reason, the last line is positive because $S(x, p)$ is increasing in p, and all quantities are positive. Thus, this proves that $TU : C_1 \to C_1^s$, completing the proof. $\square$

Chapter 2

A Quantitative Theory of Early Skills Formation and Parental Choices

2.1 Introduction

Heterogeneity at age 20 has been shown to be one of the most important determinants of lifetime inequality[1]. A significant part of such heterogeneity builds up during childhood, which is known to be a crucial phase for skills development[2]. Patterns of child care time suggest that parental investment plays an important role in shaping inequality in skills accumulation: for example, more educated parents tend to spend more time with their children in the US. Moreover, parents who perform well in cognitive tests tend to have children who perform well too[3]. Understanding the origins of such patterns is crucial for understanding inequality and evaluating the long term impact of redistributive policies. In this paper I develop a theory of parental investment in which heterogeneity in rates of returns (that is, gains in children's skills per unit of investment) provides an explanation for the cross-sectional patterns of child care time and the intergenerational correlation of test scores.

For this purpose I introduce the estimates of the technology of cognitive and noncognitive skills formation proposed by Cunha, Heckman and Schennach (2010) in an heterogeneous agents decision

[1]See Huggett, Ventura and Yaron (2011); Lee and Seshadri (2012); Guvenen and Kuruscu (2009); Keane and Wolpin (1996).

[2]The empirical evidence dates back to the Perry Preschool Project (1962) and the Coleman Report (1966); see for instance Heckman, Malofeeva, Pinto and Savelyev (2010b) and Heckman, Moon, Pinto, Savelyev and Yavitz (2010a), and also the Head Start and Early Head Start programs.

[3]See Guryan, Hurst and Kearney (2008), Ramey and Ramey (2010) on child care time differentials; Black, Devereux and Salvanes (2008), Anger and Heineck (2009) on intergenerational correlations of test scores.

theoretic model of parental investment choices and skills development. In my model, households care for their children's skills and invest time and goods in them; time trades off with work hours and leisure, while goods trade off with consumption. The main mechanism is simple: higher returns give a higher incentive to invest, hence parental choices are influenced by the structure of the technology of skills formation.

My goal is twofold. First, I want to understand the role played by parental time versus goods as investment inputs. To do so, I calibrate the model to replicate features of the US economy and to stylized facts on early skills development. I find that time with children is the most important input of parental investment, in line with the existing literature. Then I investigate whether the theory, joint with the properties of the technology, can account for noncalibrated facts observed in the data. I find that the model accounts very well for the bulk of cross-sectional variation of child care time and for correlations of test scores.

The technology of skills formation estimated by Cunha, Heckman and Schennach (2010) has a number of properties which are crucial for the success of my theory in accounting for the data. First, childrens' skills exhibit self-productivity: higher initial skills lead to higher skills later on. Also, early parental investment produces long-lasting effects because increasing skills at the beginning affects all the subsequent path of skills development. Second, during early childhood (ages 0-6) investment is more productive than during later childhood (ages 6-14). These first two properties imply that parents invest more during early childhood.

Third, higher-skilled parents are more productive in raising skillful children: this implies that higher-skilled parents have a higher incentive to invest. Also, higher-skilled parents are more likely to be college-educated: despite the fact that they earn higher wages, I find that productivity differentials can offset differences in the opportunity cost of time, providing an explanation for why college-educated parents spend more time with their children. Fourth, investment in early and later childhood are strongly complementary, which means that some amount of investment during late childhood is needed in order to sustain previously accumulated skills. All previous properties imply that parents who invest more during early childhood will tend to do so also later on, in order not to lose skills accumulated by their children: this explains why college-educated parents invest more both during early and late childhood. Cross-households productivity differentials and complementarity of early and late investment also contribute to explaining the observed correlation of cognitive skills at age 14 between mothers and children: higher incentives to invest for higher-skilled parents lead to persistence in skills inequality.

Lastly, the elasticity of substitution between inputs of skills formation varies widely across early and later childhood: that is, increasing skills during early childhood is relatively easy, while in subsequent years it becomes almost impossible due to a combination of low productivity of investment and strong complementarity between inputs.

In short, if parents invest only during early childhood and then forget about their child later on, the child's skills will diminish greatly. If instead they do not invest during early childhood, increasing the child's skills later will be extremely costly. Since households anticipate future phases of skills development, they trade off easy gains during early childhood with the fact that increasing skills "too much" requires more investment to sustain them later. Hence the effect of such property on the life cycle dynamics of parental investment is nontrivial. I find that, quantitatively, the difference in elasticities of substitution estimated by Cunha, Heckman and Schennach accounts very well for the ratio of early to late child care time observed in the data. By contrast, when this property is not accounted for, I obtain early/late ratios that are inconsistent with the empirical evidence.

The theory is succesful in accounting for several facts on child care time as observed in the 1975-1985 American Time Use Survey. In particular, the model matches well the average ratio of early (ages 0-6) to late (ages 6-14) child care time, the fact that households partly give up labor supply when a child is born, the fact that higher-educated parents invest more time in their children. As a calibration result, I find that parental time accounts for the largest share of investment, consistently with the results of Schoelmann (2012) and Del Boca et al. (2012), although goods do play a role.

The model also reproduces well the pattern of intergenerational correlations of parental and children's skills observed in the National Longitudinal Survey of the Youth 1979. I find that, to get this result, it is very important to account for the time structure of the technology, i.e. build a model that has as many time periods as the estimated technology: three periods of early childhood (0-6) and four periods of late childhood (6-14). Simplified models can explain the patterns of child care time, but fail in delivering quantitatively consistent persistence across generations[4]. The reason is that such correlations build up over time, and that many time periods are needed for parental skills to account for a larger fraction of their children's skills.

While several papers try to understand how parental choices interact with child development, returns to investment are usually very difficult to identify, because few data sets provide enough consecutive measurements of children's skills. Moreover, data on early achievement of children are often plagued

[4]For instance, Cunha, Heckman and Schennach (2010) use a two-period model with one period of early childhood and one period of later childhood to discuss the effects of different investment policies.

by measurement error issues[5]. The data of Cunha et al. allow to overcome this last problem, because they contain many measurements of children's skills. However, such data do not have information on parental time. Hence in their work parental investment is identified using different measurements[6]. My model provides a theory of how returns (as identified by the technology) affect investment decisions of households on time and goods. As a result, the patterns of child care time produced by the model are a consequence of returns to parental investment, but I emphasize that they do not simply replicate patterns of the data used to identify the technology.

Finally, I use the model to perform a policy experiment applying the 2012 German scheme of child allowances in the model. The policy could improve slightly average cognitive skills at age 14; gains in skills are higher among children from poorer families. Also, such a policy would reduce the correlation between household income and offspring's cognitive skills by 4.5 % and the intergenerational persistence of noncognitive skills by 1.5 %, while having a negligible impact on the intergenerational persistence of cognitive skills. This result is compatible with the fact that the US and Germany have very similar intergenerational correlations of cognitive test scores (Anger and Heineck (2009)), despite having very different institutional setups regarding child allowances. The intuition for my results is that, while a flat transfer can reduce relative differences in income across households by diminishing the relevance of income "per se", it cannot compensate for the fact that higher-skilled parents are more productive in raising skillful children.

The paper is organized as follows. I present a brief review of the related literature in section 2. Section 3 describes in detail the features of the technology of skills formation and the theoretical model. Section 4 describes the data used to identify the parameters and the calibration. Section 5 presents quantitative results and external validation; section 6 presents several robustness checks. Section 7 describes the results of the policy experiment. Section 8 concludes the paper.

2.2 Related Literature

The central ingredient of my model, the technology of skills formation, is taken from the work of Cunha, Heckman and Schennach (2010), who employ a novel identification strategy based on measurement-error models and estimate such technology. They find that measurement error is extremely relevant

[5]See Cunha, Heckman and Schennach (2010) on the first-order importance of measurement error in estimating skills formation technologies.

[6]For instance, questions such as "how many push/pull toys does the child own?"? See Cunha, Heckman and Schennach (2010) for details.

for measures of child development and parental investment; that cognitive and noncognitive skills, both as child's and parental endowments, contribute to each other's production; that cognitive skills can be heavily influenced in early childhood, but not in late childhood; that accounting for both cognitive and noncognitive skills production makes a big difference in the discussion of optimal policy. The work of Cunha et al. is not the only attempt to estimate technologies of skills formation; see for instance Helmers and Patnam (2011), Todd and Wolpin (2007) or Hanushek and Woessmann (2008). I choose the former technology because of its generality[7].

To the best of my knowledge, not many papers try to quantify the causes of the patterns of child care time in a structural model. Several papers build models for policy analysis using data on both time with children and cognitive achievement, for instance Bernal and Keane (2010) (2011), Del Boca, Flinn and Wiswall (2012), Griffen (2011) and Brilli (2012). I take a different route: I use external estimates of the technology of skills formation and try to understand whether differences in returns to investment can account for facts on time with children and cognitive/noncognitive test scores. The paper by Del Boca et al. is in many respects complementary to the present one; the authors estimate a structural model of cognitive skills formation and parental choices using data from the Child Development Supply of the PSID. The main message of their paper is that household choices regarding investment in children cannot be understood without also considering labor and consumption choices. While their model is built in order to understand the average effect of policies, I focus on inequality in children's skills and in parental investment that arises from the interaction of household tradeoffs and the properties of the skills formation process. Moreover, the policy results of Del Boca et al. are substantially negative, as they find that a governmental contribution to parental investment goods as large as about 25 % of median yearly household income increases latent childrens' skills by a relatively small amount over the baseline. As I show in the Policy subsection, due to big differences in the specification of the technology of skills formation I use, my policy results are substantially more encouraging in this regard.

I use the model to understand what drives the variation in time spent with children across households; a vast empirical literature studies these patterns. In particular Bianchi (2000), Guryan, Hurst and Kearney (2008) and Ramey and Ramey (2010) provide analyses of child care time in the US. I isolate two facts from this literature: parental time with children has been remarkably stable from 1960 to 1990s, and college-educated parents tend to spend more time with their children in the US[8].

[7] I also show that a technology which takes into account only cognitive skills performs poorly on several noncalibrated moments.

[8] Guryan, Hurst and Kearney (2008) show that college-educated parents spent about 35 % more time with their children in the 2000s. In a comment to the paper of Ramey and Ramey (2010), Hurst shows that such child care time differential

Finally, whether parental time or purchasable goods are more important for investment in children is an open question, and a very relevant one for my work. The empirical study of Schoellmann (2012) on Indochinese refugees suggests that goods are quantitatively unimportant in explaining differences in human capital. The model I develop predicts a greater role for parental time compared to income, consistently with his results.

2.3 The model

The model features three key ingredients: the technology of skills formation, time-allocation choices of households and ex-ante heterogeneity in skills for parents and offsprings.

Since most of the results are a direct consequence of returns to investment, the next subsection summarizes the main features of the technology of skills formation estimated by Cunha, Heckman and Schennach (2010) (CHS from now on) and the main mechanisms behind differentials in returns.

2.3.1 The technology of skills formation

CHS estimate the technology of skills formation assuming there exist two different developmental stages, $j = \{1, 2\}$, which correspond to *early childhood* (ages 0-6) and *later childhood* (ages 7-14) respectively. Early human capital of the child is assumed to be a two dimensional, time varying vector of skills; the latter are of type $k = \{C, N\}$, respectively cognitive and noncognitive skills.

Parents provide three different inputs for child development: their cognitive and noncognitive skills $s_{C,P}, s_{N,P}$, which are assumed to be time-invariant, and parental investment I_t. Parental skills are assumed to be those of the mother.

The technology of skills formation has the common Constant Elasticity of Substitution form

$$s_{k,t+1} = [\gamma_{j,k,1} s_{C,t}^{\phi_{j,k}} + \gamma_{j,k,2} s_{N,t}^{\phi_{j,k}} + \gamma_{j,k,3} I_t^{\phi_{j,k}} + \gamma_{j,k,4} s_{C,P}^{\phi_{j,k}} + \gamma_{j,k,5} s_{N,P}^{\phi_{j,k}}]^{1/\phi_{j,k}} \tag{2.1}$$

which states that next period's skills $s_{k,t+1}$ are a function of investment I_t, offspring's cognitive and noncognitive skills $\{s_{C,t}, s_{N,t}\}$ at time t and parental cognitive and noncognitive skills $\{s_{C,P}, s_{N,P}\}$. Notice that there are two stages j but many time periods t, which belong to one of the two stages. In the work of CHS, periods t are two years long and stages 1 and 2 correspond to ages 0-6 and 7-14,

was already around 11 % in 1985. Evidence on the increasing gap in parental investments between college and noncollege parents has raised concerns on whether future persistence of inequality might rise as a consequence. See for instance http://www.nytimes.com/2012/07/10/opinion/brooks-the-opportunity-gap.html?_r=3&src=me&ref=general&

respectively. All parameters $\gamma_{j,k,i}$ and $\phi_{j,k}$ vary across developmental stages $j = \{1, 2\}$ and across skills $k = \{C, N\}$. The parameter $\phi_{j,k} \in (-\infty, 1]$ is crucial, because it determines the elasticity of substitution $1/(1 - \phi_{j,k})$ between inputs.

Cunha, Heckman and Schennach estimate the technology of skills formation under a number of alternative assumptions (household-specific heterogeneity and endogeneity of investment); their findings are robust to the alternatives. I summarize below the findings that drive results in the present paper.

1. **Self-Productivity**: skills exhibit self-productivity in the sense that $\gamma_{j,C,1} > 0, \gamma_{j,N,2} > 0$ for $j = \{1, 2\}$; higher initial skills lead on average to higher skills later on. Also, early investment produces long-lasting effects because increasing skills at the beginning affects all the subsequent skill development.

2. **Cross-Productivity**: skills positively contribute to each other, in the sense that $\gamma_{j,C,2} > 0, \gamma_{j,N,1} > 0$ for $j = \{1, 2\}$. Higher cognitive skills increase noncognitive skills, and viceversa.

3. **Efficiency**: in the first stage, investment is more productive than in the second stage, for both cognitive and noncognitive skills; that is, $\gamma_{1,k,3} > \gamma_{2,k,3}$ for $k = \{C, N\}$.

4. **Complementarity**: in the first stage of cognitive skills development, the elasticity of substitution between inputs is roughly four times larger than in the second stage; this means that, during early childhood, parental investment can make up for adverse initial conditions (i.e. below-median initial cognitive endowments) and for low parental skills. During later childhood, instead, inputs become strongly complementary, so that increasing skills in this phase becomes extremely costly. Noncognitive skills, instead, exhibit roughly the same elasticity of substitution across stages.

The features of the technology, along with its estimated parametrization (available in table 2.14 in the Appendix), produce a number of derived results that give insights on how parental investment should behave if households knew the technology of skills formation. First of all, in the first stage it is easier to increase cognitive skills; the amount of investment required to increase skills by 1 % of a standard deviation is lower in the first stage with respect to the second, as figure 2.1 shows. Noncognitive skills, instead, do not exhibit such a clear pattern for the productivity of investment.

Given that returns to investment are larger in the first stage, if parents care more for cognitive rather than noncognitive skills of their offspring, we should expect investment to be higher in early childhood rather than in later childhood.

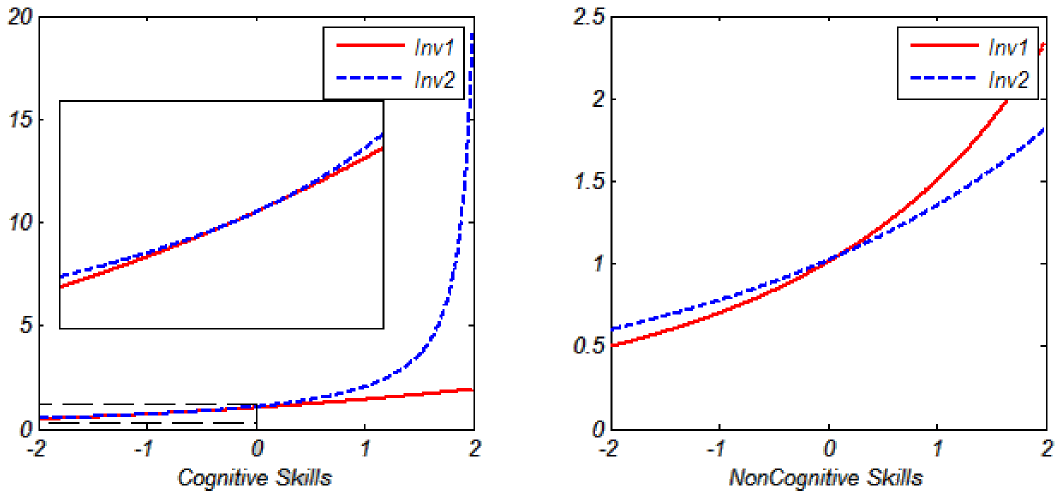

Figure 2.1: Amount of investment required to increase skills by 1 % of a standard deviation, by level of log standardized initial skills, in the first stage (red line) and second stage (blue line); parental skills fixed at the median. Graph includes magnification for lower-than-median initial cognitive skills.

Another feature of the technology is that investment in the second and first stage are strongly complementary: this happens because first stage investment enters second-stage skills production through the self-productivity of future periods' skills. Hence, the more investment is performed today, the more it is required tomorrow, even only to keep skills constant.

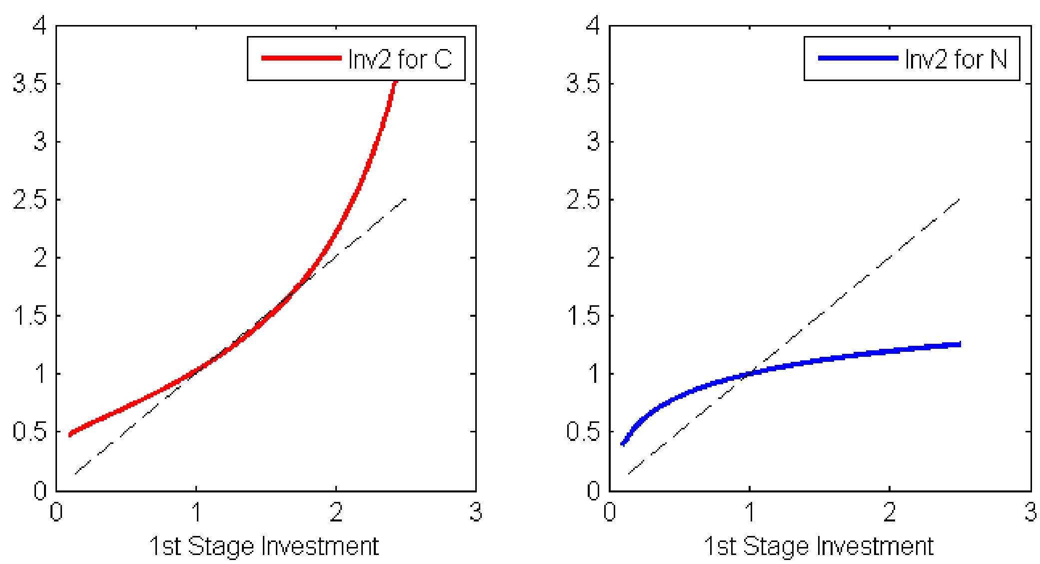

Figure 2.2: Amount of second-stage investment required to maintain skills constant, by initial investment, against 45° line; child's initial skills and parental skills fixed at the median.

Figure 2.2 shows how much investment is required in order to keep skills constant in the second stage, after investing x units in the first stage, for a median household.

The natural consequence of these two features is that we expect investment to be "smoothed" across phases, on average; moreover, household groups who invest more in the first stage will, on average, invest more also in the second stage.

The final feature I discuss here is that high-skilled parents are more productive in raising skillful children; figure 2.3 summarizes this feature of the technology.

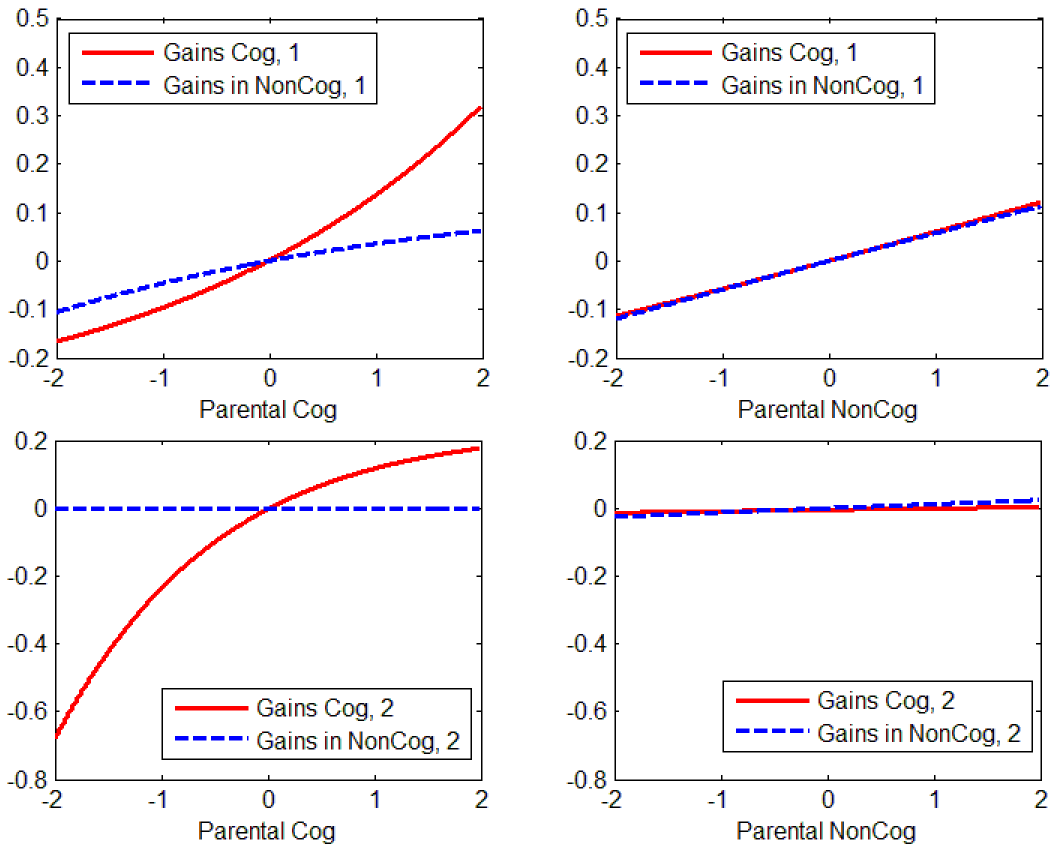

Figure 2.3: Gain in skills (as fraction of a standard deviation) by log standardized parental skills and by developmental stage; initial child's skills fixed at the median.

For instance, when a mother's cognitive skills are one standard deviation above the median, the first-stage gains in the child's cognitive skills are higher by 10 % with respect to what the median mother would produce. In general, higher parental skills yield to higher offspring's skills; and these gains are larger during early childhood than later childhood.

2.3.2 Investment in Children

The technology of skills formation estimated by Cunha, Heckman and Schennach allows to quantify returns from investment in children; in order to link these returns to the patterns of child care time, a mapping between observables and the abstract concept of "investment" is required. CHS identify investment from a large number of measurements which include whether the child has access to education goods (such as theatres, museums, musical shows, books, musical instruments), the number of specific toys children own and others. In the present paper, I make the assumption that investment at time t can be expressed as

$$I_t = A x_t^{\alpha} e_t^{1-\alpha} \tag{2.2}$$

where x_t is primary child care time spent by the household with its offspring, and e_t is the amount of goods spent as complements of child care. The complementarity of time and goods introduces a degree to which richer households are advantaged with respect to poorer households[9]; in the preferred calibration, the mechanism is shown to be important in explaining differentials in investment by education groups.

2.3.3 The Model

Finally, a decision theoretic model that embeds the two previous ingredients is developed, in order to rationalize the observed cross-sectional patterns of investment in children. The model is a parental choice model in the spirit of Becker and Tomes (1976) (1979), and is mainly based on the one which Cunha and Heckman (2007) use to rationalize the technology of skills formation. There are three main differences between this model and theirs: I include a time tradeoff, consider investment as a combination of child care time and goods and I assume parents to care for the "quality" of their child in every period.

Households face a time allocation problem and a resource allocation problem. Time is limited and must be allocated among work, time invested in children and leisure; resources come from labor income and must be allocated between consumption and goods invested in children.

2.3.4 Environment and Timing

The economy is populated by a continuum of households of measure 1, each of which has one offspring as in Cunha and Heckman (2007). Time is discrete (indexed by t) and there are $T+1$ time periods, $t = 1, 2, ...T_1, ..., T+1$ where periods $1, ..., T_1$ belong to early childhood, $T_1+1, ...T$ belong to late childhood and $T+1$ denotes the terminal period.

[9] In a robustness check, I show that if time and goods are more substitutable, the model generates some counterfactual predictions.

In the first period, parents are endowed with cognitive skills $s_{C,P}$ and noncognitive skills $s_{N,P}$, which are assumed to be time-invariant; these will be referred to as "parental skills". Every child is also endowed with initial skills $\{s_{C,1},\ s_{N,1}\}$, which will be referred to as "offspring's skills". Households are heterogeneous in initial conditions $\left\{s_{C,1}, s_{N,1}, s_{C,P}, s_{N,P}\right\}$ and, as is standard in the literature, are assumed to have full knowledge of them.

In the economy there exists only one good, which is used as the numeraire; such good can be indifferently consumed or used for investment in children.

2.3.5 Preferences and Choices

In periods $1, ..., T+1$, households decide how to allocate one unit of time into working n_t, time with their offspring x_t and leisure, how much good c_t to consume and how much to spend in goods e_t for their offspring. The household gets labor income $w(s_{C,P}, s_{N,P}, \epsilon_P^w)\ n_t$, where the wage is a function of parental skills plus additional household-specific heterogeneity ϵ_P^w, and n_t is the amount of time spent working in period t[10].

Following Restuccia and Urrutia (2004), there is no financial asset that allows redistribution of resources between time periods 1 and 2.[11] Hence the budget constraint of an household can be written as

$$c_t + e_t \leq w(s_{C,P}, s_{N,P}, \epsilon_P^w)n_t \quad \text{for } t = 1, ..., T+1 \tag{2.3}$$

The consumption good c_t gives CRRA utility $\frac{c_t^{1-\theta}}{1-\theta}$. Leisure gives utility $\frac{\zeta(1-n_t-x_t)^{1-\sigma}}{1-\sigma}$; households discount future outcomes at the common rate β.

Finally, using the terminology of Becker and Tomes (1976), at each period the "quality" of children in terms of both cognitive skills $s_{C,t}$ and noncognitive skills $s_{N,t}$ maps into parental utility $W(s_{C,t}, s_{N,t})$, weakly increasing in both arguments and strictly concave (that is, $W_{s_{C,t}} \geq 0$, $W_{s_{N,t}} \geq 0$ and the Hessian of W is negative definite for every level of skills)[12].

[10] The underlying assumption is that time of the father and of the mother are perfect substitutes. One might think that some activities can exclusively be performed by the mother, i.e. breastfeeding; however, empirical evidence suggests that although children of single mothers are at a disadvantage, such disadvantage is too small to be reconciled with strong complementarity between fathers' and mothers' time. For instance Carlson and Corcoran (2001) show that the difference between children in cognitive scores of single-parent households with intact households is statistically insignificant after controlling for income and Army Force Qualification Test score of the mother.

[11] In a sense, this is a very strong form of borrowing constraint; however, inside each period resources can be freely moved in time. The assumption is simplistic but allows to take into account long-term constraints in resources in the simplest way.

[12] Such assumption is standard in models of parental choices, see for instance Del Boca et al. (2012) and Brilli (2012).

2.3.6 Investment in children

Offspring's skills evolve according to the two-stage production function described in the first subsection[13]:

$$s_{k,t+1} = [\gamma_{j,k,1}s_{C,t}^{\phi_{j,k}} + \gamma_{j,k,2}s_{N,t}^{\phi_{j,k}} + \gamma_{j,k,3}I_t^{\phi_{j,k}} + \gamma_{j,k,4}s_{C,P}^{\phi_{j,k}} + \gamma_{j,k,5}s_{N,P}^{\phi_{j,k}}]^{1/\phi_{j,k}} exp(\eta_{j,k}) \tag{2.4}$$

for time $t = 1, ...T_1, ..., T$, stages $j = \{1, 2\}$ and skills $k = \{C, N\}$. The technology exhibits constant returns to scale, that is, $\sum_{i=1}^{5} \gamma_{j,k,i} = 1$, for $j = \{1, 2\}$ and $k = \{C, N\}$. Investment I_t is given by the combination of time x_t and goods e_t described in equation 2.2. Finally, shocks $\eta_{j,k}$ are assumed to be independently normally distributed and realize at the end of the period, hence households have to form expectations on next period's skills of the child for all possible realizations of the shocks.

2.3.7 Dynamic Problem

The state of each household at time t can be described by the current period's skills of her offspring plus the additional income variability ϵ_P^w, where the household-specific characteristics $s_{C,P}, s_{N,P}, \epsilon_P^w$ are constant while offspring's skills $s_{C,t}, s_{N,t}$ evolve over time. Hence the problem of a household in period t which belongs to developmental stage $j \in \{1, 2\}$ can be written as follows:

$$V_t(s_{C,t}, s_{N,t}, s_{C,P}, s_{N,P}, \epsilon_P^w) = \max_{c_t, e_t, n_t, x_t} \frac{c_t^{1-\theta}}{1-\theta} + \zeta\frac{(1-n_t-x_t)^{1-\sigma}}{1-\sigma} + W(s_{C,t}, s_{N,t}) + \beta\mathbb{E}\Big[V_{t+1}(s_{C,t+1}, s_{N,t+1}, s_{C,P}, s_{N,P}, \epsilon_P^w)\Big]$$

Other models assume that parents only care for the future continuation value of their children. However, in the technology of skills formation I assume, investment depreciates over time. In a robustness check, I show that if parents get utility only from the final quality of their children, this implies that investment is increasing over time, because investment in the first periods depreciates. Such a pattern is completely counterfactual, as the data show that time invested in children is decreasing in the age of the child.

[13]One potential concern is that, by assuming the functional form estimated by Cunha et al., I am already "assuming my results"; however, it is not straightforward to generate patterns of investment compatible with the stylized facts on child care time and with the correlations observed in the data. Different parametrizations of the utility and investment function yield counterfactual implications on the NLSY79/CNLSY79 and on the AHTUS data. Also, I use a simplified investment function that I compare to time use surveys, while the data with which Cunha, Heckman and Schennach estimate their technology do not include any information regarding parental time.

$$\text{s.t. } c_t + e_t \leq w(s_{C,P}, s_{N,P}, \epsilon_P^w) n_t$$
$$I_t = A x_t^{\alpha} e_t^{1-\alpha}$$
$$0 \leq n_t + x_t \leq 1, \quad n_t, x_t \geq 0$$
$$s_{C,t+1} = [\gamma_{j,C,1} s_{C,t}^{\phi_{j,C}} + \gamma_{j,C,2} s_{N,t}^{\phi_{j,C}} + \gamma_{j,C,3} I_t^{\phi_{j,C}} + \gamma_{j,C,4} s_{C,P}^{\phi_{j,C}} + \gamma_{j,C,5} s_{N,P}^{\phi_{j,C}}]^{1/\phi_{j,C}} \exp(\eta_{j,C})$$
$$s_{N,t+1} = [\gamma_{j,N,1} s_{C,t}^{\phi_{j,N}} + \gamma_{j,N,2} s_{N,t}^{\phi_{j,N}} + \gamma_{j,N,3} I_t^{\phi_{j,N}} + \gamma_{j,N,4} s_{N,P}^{\phi_{j,N}} + \gamma_{j,N,5} s_{N,P}^{\phi_{j,N}}]^{1/\phi_{j,N}} \exp(\eta_{j,N})$$
$$\eta_{j,C} \sim \mathcal{N}(0, \sigma_{\eta_{j,C}}^2), \quad \eta_{j,N} \sim \mathcal{N}(0, \sigma_{\eta_{j,N}}^2).$$

At the end of skills development, parents still get utility from the quality of their children, but cannot influence them anymore. Their maximization problem becomes

$$V_{T+1}(s_{C,T+1}, s_{N,T+1}, s_{C,P}, s_{N,P}, \epsilon_P^w) = \max_{c_{T+1}, e_{T+1}, n_{T+1}, x_{T+1}} \frac{c_{T+1}^{1-\theta}}{1-\theta} + \zeta \frac{(1 - n_{T+1} - x_t)^{1-\sigma}}{1-\sigma} + W(s_{C,T+1}, s_{N,T+1})$$

$$\text{s.t. } c_{T+1} + e_{T+1} \leq w(s_{C,P}, s_{N,P}, \epsilon_P^w) n_{T+1}$$
$$0 \leq n_{T+1} + x_{T+1} \leq 1, \quad n_{T+1}, x_{T+1} \geq 0$$

2.3.8 Economic Mechanisms and trade-offs

First order conditions imply that households trade off consumption and leisure following the equation

$$(1 - n_t - x_t) = \left(\frac{\zeta c_t^{\theta}}{w} \right)^{1/\sigma} \tag{2.5}$$

Notice that the labor choice cannot be separated from the choice of time with the offspring, so that an increase in the future value of investment in the offspring will yield, *ceteris paribus*, to a variation in labor time.

Taking derivative with respect to x_t yields

$$\zeta (1 - n_t - x_t)^{-\sigma} = \mu_t A \alpha x_t^{\alpha-1} e_t^{1-\alpha} \tag{2.6}$$

where μ_t is the multiplier associated to the investment equation at time t, which encompasses the combination of the productivity of investment and the shadow value that each agent attributes to her offspring's future skills.

Equation 2.6 states that in an interior optimum, the marginal value of leisure must equal the marginal value of investment in the offspring. Substituting equation 2.5 inside 2.6 yields

$$wc_t^{-\theta} = \mu_t A\alpha\big(\frac{e_t}{x_t}\big)^{1-\alpha}. \tag{2.7}$$

which makes clear the trade-off faced by agents between consumption and investment; working time (hence consumption) must be traded off with time invested in the offspring.

In an interior solution, investment goods are proportional to the wage w of the agent and to total time spent with the offspring x_t.

$$e_t = w\big(\frac{1-\alpha}{\alpha}\big)x_t \tag{2.8}$$

This property follows from the Cobb-Douglas form of the investment function, which gives a unique optimal ratio between the two variables. The solution is always interior in the baseline model; however, results are easily extendable to the case in which there exists a government that transfers resources to households, so that the latter might work zero hours if such transfers are large enough. If $n_t = 0$, investment goods satisfy instead the equation

$$e_t = \frac{(1-\alpha)}{\alpha}\ \zeta(1-x_t)^{-\sigma}x_t(z_t - e_t)^{\theta} \tag{2.9}$$

where z_t represents lump sum transfers from the government. Equation 2.8 is a particular case of equation 2.9, in which equation 2.5 is used to simplify the marginal utility of leisure and of consumption.

Notice that equations 2.8 and 2.9 give the solution for e_t even in the boundary case in which child care time x_t is equal to zero; this can occur if the marginal product of investment is finite for $I_t = 0$. Since the technology of skills formation is of the CES form, this happens if the complementarity parameter ϕ is lower than zero; for $\phi > 0$, the solution for x_t is always interior.

As a straightforward consequence of equations 2.2 and 2.8, the investment function in an interior solution becomes

$$I_t = Ax_t\left[w\big(\frac{1-\alpha}{\alpha}\big)\right]^{1-\alpha} \tag{2.10}$$

so that both the wage and time spent with children matter for the evolution of skills.

Clearly, in the terminal period $T+1$, since households cannot influence their child anymore, maximization implies $x_t, e_t = 0$; n_t still satisfies equation 2.5.

If the share of time in the investment function is larger than the share of goods ($\alpha > 1/2$), two main features of the model arise from the FOCs. First, households will spend more time in child care in the most productive stage; second, depending on the parameters of the utility function, they may also choose to work less hours.

Proposition 1: *Consider the multipliers μ_t associated to the constraint $I_t = Ax_t^{\alpha}e_t^{1-\alpha}$ as the function $\mu_t = \mu_t(K_t, S_t, x_t)$, where $\frac{\partial \mu_t}{\partial K_t} > 0$.*

- *Suppose that $\alpha > 1/2$; then we have that $\frac{\partial x_t}{\partial K_t} > 0$, that is, agents respond to increased productivity in investment by increasing time invested in the offspring.*

- *If $\alpha > 1/2$, preferences satisfy Balanced Growth Path ($\theta = 1$) and we have that $\sigma \in [0,1], \zeta > \frac{1-\alpha}{\alpha\sigma}$, then $\frac{\partial n_t}{\partial K_t} < 0$, that is agents respond to higher productivity in investment by decreasing hours of work.*

2.4 Data and Calibration

I choose calibration targets from the dataset of Cunha, Heckman and Schennach (2010) and from surveys of work hours and child care time carried out from 1975 to 2000. The dataset of CHS is a collection of variables regarding 2207 firstborn white children from the CNLSY/79 sample. Children in the dataset have been assessed every 2 years, along with their mothers, starting in 1986. Assessments start at birth and end at age 14; they include several measures of cognitive achievements, such as the PIAT mathematics and reading comprehension tests, and measures of noncognitive achievements such as temperamental scores. For very early ages (0-2), the best predictors of future tests are measured; for instance, for estimating cognitive skills at birth, the authors use gestation length, birth weight and motor-social development.

I obtain part of the targets from the estimation of "skills factors" from assesments of children and mothers. Following CHS, the statistical tool employed is factor analysis; the idea is that a set $[Z_1, ..Z_i, ..Z_M]$ of variables, such as tests of mathematical and reading abilities, are error-contaminated measurements of the underlying cognitive and noncognitive abilities $\{s_C, s_N\}$ of an individual. Then, each measurement i is assumed to be related to the unobservable skills of individual j at time t according to

$$Z_{i,j,t} = \alpha_{i,t} + \beta_{i,t} log(s_{C,j,t}) + \epsilon_{i,j,t} \tag{2.11}$$

so that the underlying latent variables $s_{C,j,t}, s_{N,j,t}$ can be identified from the covariance between measurements up to the normalization of one of the coefficients $\beta_{i,t}$ [14]. In this study, the latent variables are simply obtained by taking the first principal factor of several different measurements for cognitive skills and noncognitive skills, taken in the same year; the underlying identifying assumption is that, for two measurements i, j of the same child such that $i \neq j$, $\mathbb{COV}(\epsilon_{i,t}, \epsilon_{j,t}) = 0$. Cunha, Heckman and Schennach identify the factors within the estimation procedure of the technology; I choose a different strategy because of simplicity and transparency, but in principle I could use the same factors as targets for the model. Part of the correlation matrix between offspring's skills and parental skills at the end of early childhood will be taken as targets of the model. For consistency in the use of the technology, I estimate the factors following closely the choice of variables described in Cunha, Heckman and Schennach (2010).

Table 2.1 provides basic statistics on the variables in the dataset for ages 5-6 and 13-14, along with the portion of variance that can be accounted for by the corresponding factor; this is represented by the R^2 of the OLS regression in equation 2.11, where Z_i is the measurement in question. CHS report a similar statistic for their variables; the patterns I report are very close to those of CHS, which means that the factors I obtain are comparable. It is easy to see how the estimated factors account for a large fraction of the variance of the tests.

In line with Cunha, Heckman and Schennach (2010) I consider parental skills to be the mother's skills, skills of the offsprings at the end of early childhood to be the latter's skills at ages 5-6 and final skills to be offspring's skills at age 13-14. Such skills have a data counterpart in the factors calculated from tests. In the Appendix I present the estimated correlation matrix for offspring's skills and parental factors, at the end of early childhood (table 2.16) and at ages 13-14 (table 2.17); I choose part of these to be targets for calibration.

Targets and stylized facts for child care time are calculated on the dataset of Ramey and Ramey (2010), which is the merge of several surveys of time uses from 1965 to 2008. I choose to target averages from the 1975 and 1985 waves of the American Heritage Time Use Survey (AHTUS) because the dataset of Cunha et al. starts measurements of children and mothers in 1986, and child care time averages have been remarkably stable from the 70s to the 90s (see Bianchi (2000). I do not use time diaries data for 1993 because, as Ramey and Ramey notice, several researchers have doubted its comparability with other surveys (see for instance Allard et al. (2007), Robinson and Godbey (1999), Bianchi et al. (2004)).

[14] See Cunha, Heckman, Schennach (2010) for a discussion of the application of such methodology in the context of skills formation.

	Mean	Std. Dev.	Skewness	N	R^2 of factor
Child's Cognitive Factor, Age 5-6					
Peabody Picture Vocabulary	0.475	0.906	-0.103	809	31.4 %
PIAT Math	0.271	1.039	0.886	1101	37.9 %
PIAT Reading Recognition	0.246	1.015	1.466	1074	96.5 %
PIAT Reading Comprehension	0.240	0.978	1.294	1025	95.1 %
Child's Noncognitive Factor, Age 5-6					
Behavior Problem Index/ Antisocial Raw Score	0.092	0.937	-1.107	1453	55.9 %
Behavior Problem Index/ Anxiety Raw Score	-0.066	1.067	-0.820	1461	49.9 %
Behavior Problem Index/ Headstrong Raw Score	-0.098	0.996	-0.039	1462	72.3 %
Behavior Problem Index/ Hyperactive Raw Score	0.010	0.972	-0.417	1461	58.1 %
Behavior Problem Index/ Conflict Raw Score	0.064	0.905	-1.882	1463	41.1 %
Child's Cognitive Factor, Age 13-14					
PIAT Math	0.424	0.921	-0.220	1063	64.5 %
PIAT Reading Recognition	0.336	0.876	-0.639	1064	78.8 %
PIAT Reading Comprehension	0.427	0.937	-0.270	1056	72.4 %
Child's Noncognitive Factor, Age 13-14					
Behavior Problem Index/ Antisocial Raw Score	0.117	0.971	-1.148	1125	63.5 %
Behavior Problem Index/ Anxiety Raw Score	-0.088	1.053	-0.595	1138	64.8 %
Behavior Problem Index/ Headstrong Raw Score	-0.07	0.998	0.002	1143	68.3 %
Behavior Problem Index/ Hyperactive Raw Score	0.044	0.974	-0.715	1138	59.3 %
Behavior Problem Index/ Conflict Raw Score	-0.024	1.033	-1.577	1142	52.4 %
MOTHER's Cognitive Factor					
Mom's Arithmetic Reasoning Test Score	0.172	0.933	0.168	2207	83.7 %
Mom's Word Knowledge Test Score	0.302	0.822	-0.836	2207	70.9 %
Mom's Paragraph Composition Test Score	0.377	0.827	-1.121	2207	66.0 %
Mom's Numerical Operations Test Score	0.343	0.875	-0.469	2207	54.7 %
Mom's Coding Speed Test Score	0.468	0.879	-0.445	2207	41.1 %
Mom's Mathematical Knowledge Test Score	0.185	0.972	0.269	2207	77.4 %
MOTHER's NonCognitive Factor					
Mom's Self-Esteem: "I am a person of worth"	3.534	0.516	-0.343	2207	43.1 %
Mom's Self-Esteem: "I have good qualities"	3.382	0.531	0.025	2207	48.5 %
Mom's Self-Esteem: "I am a failure"	3.477	0.580	-0.649	2207	52.9 %
Mom's Self-Esteem: "I am as capable as others"	3.326	0.549	-0.217	2207	36.7 %
Mom's Self-Esteem: "I have nothing to be proud of"	3.480	0.625	-1.082	2207	46.0 %
Mom's Self-Esteem: "I have a positive attitude"	3.200	0.576	-0.250	2207	51.6 %
Mom's Self-Esteem: "I wish I had more self-respect"	2.876	0.787	-.206	2207	38.2 %
Mom's Self-Esteem: "I feel useless at times"	2.650	0.774	0.300	2207	32.5 %
Mom's Self-Esteem: "I sometimes think I am no good"	3.005	0.808	-0.298	2207	41.9 %
Mom's Rotter Score: "I have no control"	2.897	1.156	-0.600	2207	5.5 %
Mom's Rotter Score: "I make no plans for the future"	2.543	1.159	-0.002	2207	8.1 %
Mom's Rotter Score: "Luck is big factor in life"	3.154	0.974	-1.107	2207	4.5 %
Mom's Rotter Score: "Luck plays big role in my life"	2.426	1.144	-0.025	2207	2.5 %

Table 2.1: Summary Statistics of Variables used to identify latent Cognitive and Noncognitive Factors. Source: data extract from NLSY79/CNSLY79 by Cunha, Heckman and Schennach (2010).

I also do not target later averages because, as Ramey and Ramey (2010) and Aguiar and Hurst (2007) show, child care time from 1995 onwards increases dramatically for all groups, and in particular for college-educated parents. Hence I compute 1975-1985 averages for married working fathers and mothers by education and by developmental stage of their youngest offspring; I follow the authors' definition of child care time, considering it as the sum of primary, educational and recreational child care time. My results are comparable to those of Hurst's comment to Ramey and Ramey; more details on how the averages have been calculated can be found in the Appendix.

2.4.1 Calibration

Two of the main ingredients of the model are taken from the paper of Cunha, Heckman and Schennach (2010): the parametrization of the technology of skills formation and the initial distribution of skills at birth. Parameters of the technology are reported in table 2.14, including the variance of the shocks to skills; the authors estimate several versions of the technology under different sets of assumptions, such as the existence of unobserved heterogeneity across households and endogeneity of investment; I choose the latter estimation as it already accounts for endogeneity of the investment function, making it more suitable for inclusion in a decision theoretic model.

In line with CHS, cognitive and noncognitive skills at birth and parental skills are assumed to be jointly lognormally distributed with mean zero and covariance matrix Σ. The parametrization of the covariance matrix is taken from the appendix of the paper of CHS and available in table 2.13 in the Appendix to this paper, along with the correlation matrix to allow an easier interpretation.

The technology has been estimated on two years intervals, hence I set the time span of the model so that one period corresponds to two years. Periods 1,2,3 correspond to early childhood, from when a child is born to when he is 6; periods 4,5,6,7 correspond to late childhood so that skills development is assumed to end at age 14, and period 8 corresponds to the terminal period.

The discount factor β is set to 0.92, which is equivalent to 0.96 at the yearly level, a standard value in the macroeconomics literature.

Following Osuna and Rios Rull (2003), I set the time endowment of households to be 200 hours per week, excluding sleep and personal maintenance.

The baseline value of risk aversion θ is set to 1, which implies that utility of consumption is given by $\log(c_t)$. As Erosa, Fuster and Kambourov (2012) argue, values of θ different from 1 imply that different hours of work across lifetime income groups, while the data show that individuals with different levels of lifetime income tend to work roughly the same amount of hours. I perform sensitivity analysis with the values of 0.8 and 1.2.

The curvature of leisure σ (at the micro level) is relevant for the analysis as it determines the substitution between investment in children and leisure. I set the baseline value in order to obtain a plausible elasticity of labor supply at the household level; a wide literature attempts to estimate such elasticity. For instance Pistaferri (2003) argues for a value near 0.7, while Erosa, Fuster and Kambourov (2012) use a value of about 0.6. Keane (2011) also shows that estimates of labor supply elasticity vary widely in the literature, and the average of the estimates is around 0.85. I target the value of 0.6, which lies in the upper range of the micro estimates of the early literature, and well within the range surveyed by Keane (2011). The baseline value of σ is set to 3.5, which implies a Frisch labor supply elasticity $\eta^{\lambda} = 0.58$ in the first stage for the median household, 0.53 in the second stage; I perform several robustness checks for different values. I use as a reference point the median household because the model generates a distribution of elasticities of labor supply, due to the fact that households choose jointly different levels of child care time and of labor supply.

The parameter ζ targets the average hours of work of married households aged 25-44. A well-known stylized fact is that per-capita hours of work in the US have not moved much in the last 50 years; As the average hours per week in 1980 were around 42 for married men and 26 for married women in the 25-34 and 35-44 years old age groups (without children or with a child older than 6), this gives 68 hours per household which means $68/200 = 34$ % of available time per week[15].

The share of time α in the investment function implies a larger role for time and a smaller role for income in explaining the variation of final offspring's skills. Hence α is chosen as to match the correlation between final cognitive skills of the offsprings and the residual of a Mincer regression of household income on parental skills, as observed in the NLSY/79. the rationale for this choice is that I want to isolate the effect of additional income from the effect of parental skills; the latter have a positive effect both on parental income and on the productivity of investment.

The scale A of the investment function implies a normalization of final skills; the chosen normalization is that the logarithm of final cognitive skills generated by the model has mean zero, i.e. the average skills of mothers and offsprings are equal at the end of the developmental process. One possible concern is that the psychometric literature reports that several countries have experienced massive increases (up to 1.2 standard deviations) in cognitive test scores in the last 50 years, the so called "Flynn effect" (see

[15]These numbers come from McGrattan and Rogerson (2004), which give data on hours worked of individuals from 1950 to 2000.

Flynn (2009)). However, whether this represents a generalized increase in cognitive abilities or, instead, an increase in the ability to perform well in tests, is a matter of debate. To answer this possible critique, I perform a robustness check in which the scale A targets another possible normalization. Results do not change very much depending on the chosen normalization, although taking into account the Flynn Effect yields results that are slightly further from the data.

Income at the household level depends on parental skills and on the labor supply choice; while the dataset provides the former, the latter are unobserved. I use indirect inference to make the income process generated by the model consistent with the NLSY/79 data. I first run a Mincer regression between household log income and parental cognitive and noncognitive factors, uncovering Mincer returns for the two skills[16];

$$\log Y_t = const + \beta_C s_{C,P} + \beta_N s_{N,P} + \Gamma controls + \epsilon_t \tag{2.12}$$

then, returns to skills for wages in the model are calibrated so that, when performing an analogous Mincer regression on model-generated data, the estimated coefficients match the data counterparts β_1 and β_2.

Controls include a full set of year dummies and a cubic polynomial in age of the mother. To make the coefficients of the regression consistent with the scale of parental skills in the model, factors estimated from the data are rescaled using the covariance matrix in table 2.13. Moreover, as the model has 2-years time periods, the regression is performed on the sum of 2 years of log income[17], which allows to get a better fit by diminishing the amount of income variance due to idiosyncratic shocks. Table 2.2 reports results from the estimation by random effects on the dataset.

Household-specific heterogeneity in wages ϵ_P^w is assumed to be normally distributed, independent of skills and it is calibrated to have the same variance as that observed in the residual of the Mincer regression ϵ_t estimated above.

Finally, the choice of the utility function W is somewhat more difficult, as many adult outcomes may contribute to the value parents attribute to investment in the offspring; the chosen functional form for utility given by "quality" of children is

[16] I use only the mother's factors because the dataset does not provide information on the father's skills. However, using only mother's skills allows to account for the correlation between fathers and mothers while economizing state variables in the model.

[17] The concept resembles the idea of "permanent income"; in order to get better estimates, one could sum income of more years. However, the NLSY data does not provide enough consecutive observations of income, a shortcoming that would make the estimation sample smaller and the estimates less precise.

Dep. Variable: Log Household Income	
	Mincer Equation
Norm. Mom Cognitive Factor	0.334***
	(0.03)
Norm. Mom NonCognitive Factor	0.249***
	(0.09)
Constant	1.334***
	(0.18)
R^2(within)	0.022
R^2(between)	0.211
R^2(overall)	0.177
N	
significance levels: * = 0.1, ** = 0.01 *** = 0.001	

Table 2.2: Mincer regression: Log Household Income in the NLSY/79 as a function of Log Cognitive and Noncognitive Factors of the mother; years 1988-2002, household with child older than 6 present. Controls include a full set of year dummies and a cubic polynomial in age (omitted). Source: data extract from NLSY79/CNSLY79 by Cunha, Heckman and Schennach (2010).

$$W(s_{C,t}, s_{N,t}) = \chi \frac{\left(s_{C,t}^{\psi} s_{N,t}^{1-\psi}\right)^{1-\xi}}{1-\xi}$$

χ gives the relative importance of the future of the offspring with respect to consumption of the family and leisure. The higher χ, the higher will be the value of the offspring's skills for households, the higher will be investment. Hence, this parameter targets the average amount of time parents spent in early primary child care in 1975-1985, from when the offspring is born to when he/she is five years old; I target the average of early child care time as observed in the American Heritage Time Use Survey of 1975-1985 for married parents, which is around 9 % of the available endowment[18].

ψ determines the relative importance of cognitive skills w.r.t. noncognitive skills and ξ encompasses the risk aversion in the future of the offspring. The two parameters are identified by targeting the pattern of correlations between parental skills and offspring's skills observed in the data. As skills evolve, such correlations change over time. First, a lower risk aversion implies that high-skilled parents have stronger incentives to invest, because marginal returns are higher than those of consumption; hence, a lower risk aversion increases the intergenerational persistence of skills. Following this intuition, I use ξ to match

[18]The number is calculated by taking the sum of average hours per week spent by mothers and of average hours per week by fathers, and dividing it by the stock of available hours per week, assumed to be 200 hours. This target may be very different depending on the definition of child care time / time with children that are considered as investment; see for instance Ramey and Ramey (2010), Bianchi (2000), Sandberg and Hofferth (2001), Del Boca, Flinn and Wiswall (2012)

the correlation between parental skills and offspring's skills at the end of early childhood. Finally, the relative weight of cognitive skills ψ implies that households will invest more in the first stage. From the parametrization of the technology (table 2.14), it can be seen that in the first stage noncognitive skills contribute to the production of cognitive skills ($\gamma_{1,C,2} = 0.062$); if investment in this stage is increased, this reduces the relative relevance of noncognitive skills for the production of cognitive skills. Hence, the observed correlation between cognitive and noncognitive skills is negatively affected by ψ: then I use this parameter to target this last correlation at the end of early childhood.

Tables 2.3 (externally set parameters) and 2.4 (endogenously determined) summarize the proposed calibration; the model has no trouble in matching the targets very closely.

Parameter	Value
Technology of Skills Formation	Cunha, Heckman, Schennach (2010) (see table 2.14 in the Appendix)
Covariance Matrix of Initial Conditions	Cunha, Heckman, Schennach (2010) (see table 2.13 in the Appendix)
Duration of One Period	2 years
β	0.92
θ	1
σ	3.5

Table 2.3: Calibration of parameters/functional forms set exogenously.

The calibration shows first of all that goods by themselves are relatively unimportant for child development: the calibrated value of $\alpha = 0.77$ suggests that parental time plays the central role, and that the pattern of investment is mainly explained by the pattern of child care time.

Second, $\psi = 0.60$ suggests that cognitive skills are relatively more important to households than noncognitive skills; as I will show, this implies that early childhood investment is higher than later investment. As the productivity of investment for cognitive skills is higher during early childhood, households concentrate their efforts there.

Finally, $\xi = 0.65$ suggests that risk aversion in the "quality" of children is lower than risk aversion in consumption; in fact, sensitivity analysis on the risk aversion for consumption θ shows that calibrated values of ξ are always lower than θ.

Parameter	Value	Target	Data	Model
Preferences				
ζ	0.63	Hours worked	0.34[a]	0.34
χ	0.65	Avg. Time in Child Care when child < 6	0.091[b]	0.092
ψ	0.60	Correlation($s_{C,4}, s_{N,4}$)	0.16[c]	0.16
ξ	0.65	Correlation($s_{C,4}, s_{C,P}$)	0.28[c]	0.28
Income Equation				
$\beta_{C,\text{model}}$	0.35	Mincer returns to $s_{C,P}$	0.33	0.34
$\beta_{N,\text{model}}$	0.26	Mincer returns to $s_{N,P}$	0.25	0.25
Variance(ϵ_P^w)	0.48	Var. of Mincer Residuals	0.48	0.48
Investment				
A	17.1	Mean($s_{C,3}$) $= 0$		
α	0.77	Correlation($s_{C,T+1}, \epsilon_P^w$)	0.09[c]	0.09

[a] See McGrattan and Rogerson (2004).
[b] Source: own calculations on 1975-1985 AHTUS.
[c] Source: own calculations on CNLSY/79.

Table 2.4: Calibration of parameters endogenously determined; targets, data moments and simulated moments.

2.5 Results and Discussion

I solve the model and explore how the model performs in explaining non-calibrated stylized facts of child care time and intergenerational persistence. The model presents some computational challenges, as there are 5 continuous state variables and the dynamic problem changes at all periods due to finite life and different technologies for different phases of childhood[19]. I use polynomial approximation of the value function to solve the model. For this particular application, polynomial approximation gives very precise and reliable results; I discuss in detail the algorithm in the Appendix.

In the first subsection I focus on the behavior of child care time and labor supply over the life cycle, and on the college vs noncollege differential in child care time. In the second part, I compare the intergenerational correlations of test scores generated by the model with those in the NLSY/79, and with those that have been estimated in the literature.

[19]In principle I could discretize state variables and rely on interpolation techniques between grid nodes. However, this approach requires an enormous amount of computational time, even with very rough approximations. Moreover, I find it to yield very high approximation errors.

2.5.1 Investment and Child Care Time

First I look at how time with children varies across parental education and across developmental stages. It is well known that parental time in child care is higher when a child is ages 0 to 5. Moreover, Ramey and Ramey (2010) and Guryan, Hurst and Kearney (2008) show that higher educated parents spend more time with their children, especially during early childhood. This was already true in 1985 (see the comment of Hurst to the paper of Ramey and Ramey), but is particularly relevant from 1995 onwards; nowadays child care time by college-educated parents is about 35 % higher than that of noncollege-educated. Table 2.6 reports a comparison of key moments generated by the model with stylized facts observed in the data for 1975 and 1985; unconditional means are calculated on the data of Ramey and Ramey (2010) for parents aged 25-44. Unfortunately the AHTUS does not have information on the extent of assortative mating, i.e. when it happens that college-educated mothers are married to college-educated fathers. I partially overcome this problem by constructing two limiting cases: the case in which college-educated women are always married to college-educated men, and the case in which college-educated women are married randomly. For simplicity, these will be referred to as the "college-mating" case and the "random-mating" case. The "true" averages of household child care time by education are likely to lie between these two extrema; for instance Schwartz and Mare (2005) report that in the 1960s it was 3 times more likely to marry someone of a similar level of education than someone with a different one, and that this figure has increased steadily until the 2000s.

In order to generate college/noncollege differentials in the model, I first estimate a probit model using the dataset of CHS (2010). Then, I use the estimated probit to partition households in the model economy into those with a college-educated mother and households with a less-than-college-educated mother.

Table 2.6 summarizes the results. The model overestimates by 12 % the ratio of average early to late child care time; in the data, parents of children under 5 years old spend in early child care about 2.5 times the amount of time they put in later years, and the model replicates quite well this pattern. The model also predicts that child care time is slightly smoother for college-educated than for non-college, i.e. it has a lower early/late ratio, consistently with the data. Also, the model replicates very well the decline in work hours among parents of very young children.

The reason behind the large difference in early and late child care time is that households care for the cognitive development of their offsprings; since cognitive skills can be boosted primarily during early childhood, due to a combination of higher productivity of investment and higher elasticity of substitution

Dep. Variable: College Prob.	
Norm. Mom Cog Fac	1.360***
	(0.09)
Norm. Mom NonCog Fac	0.217
	(0.22)
constant	-0.373***
	(0.05)
Pseudo R^2	0.27
N	915
significance levels: * = 0.1, ** = 0.01 *** = 0.001	

Table 2.5: Probit model: being a college-educated mother as a function of her Cognitive and Noncognitive skills; sample includes all mothers aged > 24. Source: data extract from NLSY79/CNSLY79 by Cunha, Heckman and Schennach (2010).

between inputs, households direct most of their efforts during this stage. Moreover, the high elasticity of substitution makes returns to investment high for all households, so that even low-skilled parents prefer to invest in early childhood rather than later. However, in the later phase inputs are strongly complementary: this gives rise to two mechanisms. First, households anticipate that they will have to sustain the skills of their children also in the later stage, when it will be unlikely and costly to increase them. As a result, they invest more "smoothly" than what they would do if elasticities were the same. In counterfactual exercises in which elasticities are held the same in the two stages, the ratio of early to late child care time goes from 4 to 12 depending on the value of such elasticities. The second mechanism is that, since inputs are strongly complementary, low-skilled households invest less smoothly because they anticipate that they will have a lower productivity in the second stage; then, they prefer to invest strongly in the first stage, and to "give up" afterwards. This generates the observed difference in ratios early/late child care time across education groups, and a sizeable correlation between offspring's skills and parental skills at age 14, as we will see in next section.

Results also show that the model is able to explain 60 % of the differential in early child care time between college-educated and noncollege-educated, in the "college-mating" case-scenario. The late childhood differential in the data is more robust to the two different scenarios than the early one, and the model generates 89 % of the college-mating case average.

Tables 2.7 and 2.8 provide a summary of investment across quantiles of skills and income; such differences shed light on what underlies the college vs noncollege differential and why noncollege-educated

Parents with same education	Early Time (model)	Early Time (1975, 1985)	Late Time (model)	Late Time (1975, 1985)	Early/Late (model)	Early/Late (1975, 1985)
Mean	9.2 % (calibr.)	9.1 % (0.3 %)	3.3 %	3.6 % (0.1 %)	281 %	251 %
Mean (College)	9.6 %	9.9 % (0.5 %)	3.5 %	4.0 % (0.3 %)	273 %	244 %
Mean (NonCollege)	8.9%	8.8 % (0.3 %)	3.1 %	3.5 % (0.1 %)	287 %	247 %
Δ College/NonCollege	7.4 %	+12.6 %	12.7 %	+14.3 %	-4.7 %	-1.2 %
Educated Mother, all fathers	Early Time (model)	Early Time (1975, 1985)	Late Time (model)	Late Time (1975,1985)	Early/Late (model)	Early/Late (1975,1985)
Mean (College)	9.6 %	9.2 % (0.5 %)	3.5 %	3.8 % (0.3 %)	273 %	242 %
Mean (NonCollege)	8.9 %	9.1 % (0.3 %)	3.1 %	3.6 % (0.1 %)	287 %	252 %
Δ College/NonCollege	7.4 %	+1.8 %	12.7 %	+6.1 %	-4.7 %	-4 %
	Early Work (model)	Early Work (1980, 1990)	Late Work (model)	Late Work (1980,1990)	Early/Late (model)	Early/Late (1980,1990)
Mean	30 %	31 %	34 % (calibr.)	34 %	89.3 %	91.2 %

Table 2.6: Summary statistics for Time invested in children generated by the model and data equivalents; data are the author's calculations on the 1975-1985 American Heritage Time Use Survey, on married parents aged 25-44 of children aged 0-4 for early time and 5-17 for late time. Numbers are obtained by summing average primary child care time of mothers and average time of fathers, and dividing by the assumed time endowment of 200 hours. For "parents with same education" I sum the unconditional mean of mothers and fathers with same education; for the second case, I sum the average of mothers by education with the overall unconditional mean of fathers. All observations are weighted as recommended by the AHTUS; standard errors are reported in parentheses. Work hours data come from McGrattan and Rogerson (2004) and are calculated as the sum of average working hours for married males plus average working hours for married females.

spend relatively more time during early childhood.

Two main mechanisms drive investment; higher parental skills grant higher productivity and higher income. Goods are complementary to time: hence higher income families have a comparative advantage for investing in children. Since marginal returns to skills are higher than those to consumption, college-educated households choose to invest more time in their children, despite their higher opportunity cost.

Notice also that even higher income parents invest more time in their children, even after controlling for skills; the latter comparison highlights the role of complementarity alone, separated by the effect of skills.

	D1	Q1	Q2	Q3	Q4
Investment (time and goods)					
by Cog Skill	98.8	100.0	101.3	101.8	103.0
by NCog Skill	87.4	91.4	98.8	103.9	112.1
by ParCog Skill	78.4	84.8	97.7	106.1	117.5
by ParNCog Skill	90.9	93.4	99.9	102.7	109.9
by Income	72.8	79.7	94.4	106.2	125.8
by Inc, Q2 Cog	79.0	84.4	95.7	104.7	119.2
Time Investment					
by Cog Skill	101.9	101.6	100.3	99.2	97.6
by NCog Skill	88.9	92.0	97.8	101.7	107.2
by ParCog Skill	86.5	90.8	98.4	102.6	106.8
by ParNCog Skill	94.2	95.6	98.8	100.2	103.9
by Income	90.7	93.2	98.0	101.4	106.0
by Inc, Q2 Cog	95.9	96.9	99.0	100.7	102.8

Table 2.7: Simulated Data: Total **early investment** (time and goods) and Time Invested as percentage of the unconditional median; by first decile and quartiles of parental Cognitive Skills, parental NonCognitive Skills, parental Income and parental Income inside second quartile of Cognitive Skills.

	D1	Q1	Q2	Q3	Q4
Investment (time and goods)					
by Cog Skill	68.9	76.7	93.6	108.1	135.3
by NCog Skill	77.1	83.6	97.3	107.6	125.1
by ParCog Skill	73.1	81.3	97.8	109.2	125.4
by ParNCog Skill	87.6	91.2	100.9	105.0	116.4
by Income	75.4	82.2	96.7	108.1	126.7
by Inc, Q2 Cog	84.2	88.6	97.5	104.5	115.6
Time Investment					
by Cog Skill	74.1	80.3	94.3	106.5	127.1
by NCog Skill	79.7	85.3	97.1	106.0	119.9
by ParCog Skill	81.3	87.7	99.3	106.4	114.8
by ParNCog Skill	91.3	93.8	100.3	103.1	110.7
by Income	94.4	96.6	100.8	103.6	107.2
by Inc, Q2 Cog	102.4	101.9	101.1	100.7	99.9

Table 2.8: Simulated Data: Total **late** investment (time and goods) and Time Invested as percentage of the unconditional median; by first decile and quartiles of parental Cognitive Skills, parental NonCognitive Skills, parental Income and parental Income inside second quartile of Cognitive Skills.

2.5.2 Intergenerational Correlations

The model provides a framework to understand the persistence of cognitive and noncognitive achievement, and of income as caused by the former. In this subsection I look at correlations between parental skills and children's skills, and at their model counterparts. This is both a result and a consistency check: the fact that the model replicates well the pattern of correlations means that simulated investment is consistent with the data.

Table 2.9 reports simulated correlations between parental skills and their offspring's skills, as observed in the model and as calculated using factor analysis in the dataset of Cunha et al. I also compute the "implied" intergenerational persistence of income by calculating what would be the future wage implied by the final skills of the offspring, using the Mincer coefficients for males and females calculated by Heckman, Stixrud and Urzua (2006) on the NLSY/79.

The model performs quite well on almost all correlations, in the sense that they reside within 95 % confidence bands of the data equivalents, with one exception. One of the most interesting results is that the model can account very precisely for the observed correlation between cognitive tests of the mother and cognitive tests of her offspring at age 14 (0.45 in the model against 0.43 in the data). A recent strand of the literature looks at the intergenerational correlation of IQ scores: for instance Black, Devereux and Salvanes (2008) find that such correlation is around 0.32 between fathers and sons using Norway military data. Anger and Heineck (2009) find higher numbers for Germany: the father-son correlation is around 0.38 and the mother-daughter one is 0.48. The model generates a correlation that is fully compatible with these findings. Also, the model predicts quite well the correlation between both cognitive and noncognitive skills and household income, which shows up already at the end of early childhood, and is even reinforced at age 14. The model predicts a mother-offspring correlation of noncognitive skills at age 14 that is 30 % higher than in the data (0.20 against 0.16), but still inside the 95 % confidence bands.

The implied intergenerational correlation of earnings is 0.08 for males and 0.09 for females, and is calculated adjusting for the variance of income unexplained by skills[20]; Jäntti et al. (2006) show that

[20] The number is calculated as follows. Final skills of the offspring translate into future "average" income $\alpha+\beta_C s_{C,T+1}+\beta_N s_{N,T+1}$, where the Mincer coefficients for males and females are taken from Heckman, Stixrud and Urzua (2006), and skills are standardized to have variance = 1. Given the R^2 of this regression, it follows that the true variance of income is $1/R^2$ times the variance explained by skills alone, which implies that the standard deviation is $\sqrt{1/R^2}$ times the one implied by skills alone. Then I compute

$$\rho(\text{Offspring's Income}, \text{HH Income}) = \rho(\alpha+\beta_C s_{C,T+1}+\beta_N s_{N,T+1}, \text{HH Income})\sqrt{R^2}$$

where $R^2 = 0.099$ for males and $R^2 = 0.128$ as in Heckman, Stixrud and Urzua, which gives the numbers shown in the

	Data	Model
End of early childhood		
$\rho(s_{C,4}, s_{C,P})$	0.28	0.28
	(0.20 - 0.36)	(calibr.)
$\rho(s_{C,4}, s_{N,P})$	0.15	0.20
	(0.06 - 0.24)	
$\rho(s_{N,4}, s_{C,P})$	0.21	0.19
	(0.16-0.26)	
$\rho(s_{N,4}, s_{N,P})$	0.20	0.20
	(0.15 - 0.25)	
$\rho(s_{C,4}, s_{N,4})$	0.16	0.16
	(0.07 - 0.24)	(calibr.)
$\rho(s_{C,4}, \text{HH Income})$	0.19	0.27
	(0.06 - 0.31)	
$\rho(s_{N,4}, \text{HH Income})$	0.20	0.16
	(0.13 - 0.26)	
End of late childhood		
$\rho(s_{C,T+1}, s_{C,P})$	0.43	0.45
	(0.38 - 0.48)	
$\rho(s_{C,T+1}, s_{N,P})$	0.18	0.21
	(0.12 - 0.24)	
$\rho(s_{N,T+1}, s_{C,P})$	0.15	**0.21**
	(0.07 - 0.19)	
$\rho(s_{N,T+1}, s_{N,P})$	0.14	0.19
	(0.09 - 0.20)	
$\rho(s_{C,T+1}, s_{N,T+1})$	0.25	0.27
	(0.19 - 0.31)	
$\rho(s_{C,T+1}, \text{HH Income})$	0.30	0.32
	(0.24 - 0.36)	
$\rho(s_{N,T+1}, \text{HH Income})$	0.22	0.20
	(0.16 - 0.28)	
$\rho(\text{Offspring's Wage age 30, HH Income})$ (males)	0.36	0.08
$\rho(\text{Offspring's Wage age 30, HH Income})$ (females)	0.16	0.09

Table 2.9: Correlations between offspring's achievements and parental skills; estimates in the data have 95 % confidence bands in parentheses. Simulated correlations outside confidence bands are displayed in bold. Correlations of data factors are computed using results from the data of Cunha, Heckman and Schennach (2010).

in the US in 1980 the intergenerational correlation of the offspring's earnings with family income was about 0.36 for males and 0.16 for females. Couch and Dunn (1997) report lower numbers, although their

table. Clearly, this approach does not account for changes in prices of skills that are likely to have occurred in later years.

results depend on the year of the survey and on the inclusion or exclusion of years in which zero earnings were reported. I compare simulation results with the higher numbers reported by Jäntti: the model can explain 22 % of the persistence of income for males and 56 % for females. This is not surprising as the variance of wages that can be explained by skills at age 14 is at most 13 %; one possible reason is that parental income plays a significant role also for later human capital development, i.e. for college attendance or job search choices.

2.6 Robustness Checks

I present here a discussion of the key assumptions and of the performance of the model under alternatives; most of the main results of robustness checks are presented in table 2.18 in the Appendix.

2.6.1 A two-periods model

One might think that, once the technology has been identified for early childhood and late childhood, most results can be obtained also with a simpler two-periods model in which there is one period of early childhood, $t = 1$ and one period of late childhood $t = 2$, in the spirit of the policy experiment of Cunha, Heckman and Schennach (2010). However, given the technology of skills formation, having many periods of child development reinforces the correlation between parents and children every period. To see why, consider a very simple model of child development, in which future skills only depend on the child's own skills and parental skills in a linear fashion:

$$s_{t+1} = \gamma s_t + (1 - \gamma) s_P$$

where $\gamma \in (0, 1)$. It follows that at period $t + k$,

$$s_{t+k} = \gamma^k s_t + (1 - \gamma^k) s_P$$

so that parental skills account for a larger fraction of the child's skills at $t + k, \forall k > 1$.

Simulation results suggest that a two-periods-only model can roughly account for the patterns of child care time, but fails in delivering intergenerational correlations that are quantitatively consistent with the NLSY/79, because such correlations require more periods to build up.

2.6.2 The role of preferences

In order to study whether most of the results are a consequence of assumptions on the utility function, I compute a version of the model in which parents care for the skills of their children only in the terminal period. While such a model can be successfully calibrated to the data and is still capable of delivering quantitatively consistent intergenerational correlations of skills, many other of its implications are highly counterfactual. Due to depreciation of skills in the technology, which is particularly important in early childhood, it predicts an increasing pattern of child care time over the life cycle, which is completely inconsistent with the data. Also, it predicts that hours of work rise during early childhood, too low a final correlation between cognitive and noncognitive skills and college/noncollege differentials in child care time that are too small compared to those in the data.

2.6.3 Equal elasticities across stages

I produce a counterfactual simulation in which elasticities of substitution are equalized to 1 in all stages for all skills; that is, $\phi_{1,C} = \phi_{2,C} = \phi_{1,N} = \phi_{2,N} = 0$, so that the technology becomes of the Cobb-Douglas form. Detailed results are reported in table 2.18 in the Appendix. First of all, in order to match the intergenerational correlation at the end of early childhood, the model requires a very high weight of cognitive skills w.r.t. noncognitive ($\psi = 1$), and a low risk aversion in parental utility from offspring's skills ($\xi = 0.47$), as opposed to the more balanced results of the standard technology ($\psi = 0.60, \xi = 0.65$). As a consequence, the equal-elasticities model predicts that households invest over 6 times more in early childhood than in later childhood, which is inconsistent with the AHTUS data. However, the model is still capable of generating differentials in child care time between college-educated and noncollege-educated that are compatible with those observed in time diaries. Finally, such model performs relatively well on most correlations, but undershoots the intergenerational persistence of cognitive skills by 20 %.

2.6.4 A Model with Cognitive Skills only

CHS show that the two-skills assumption is extremely relevant for policy analysis; when they estimate a cognitive-skills-only production function, policy prescriptions move investment from disadvantaged children to advantaged ones. To investigate how a one-skill model would perform on the data, I calibrate a version of the model in which the two-skills production function is replaced with the one-skill production function estimated by CHS in the Appendix to their paper (shown in table 2.15. Results are available upon request.

The one-skill model presents two main quantitative issues; first, it is impossible to match the observed intergenerational correlation of cognitive skills at the end of early childhood, which the model grossly overestimates (0.62 against 0.28). As a result, the model predicts an intergenerational correlation at offspring's age 14 of 0.92, which is inconsistent with any study on such correlations. Second, the model cannot generate a large early/late child care time ratio, and predicts that college-educated parents spend *less* time than noncollege ones with their children: such predictions are inconsistent with the data on time uses.

2.6.5 Sensitivity Analysis

I explore whether some of the externally set parameters are playing a large role in shaping the results; for this purpose, I recalibrate the model modifying risk aversion in consumption θ and curvature of leisure σ and see how the main results change. Also, I take into account the possibility that skills of offsprings at age 14 are, on average, larger than those of their mothers (the "Flynn effect") by recalibrating the model in such a way that final cognitive skills increase by 0.5 standard deviations in a generation (which would be roughly compatible with an increase of 1 standard deviation over 50 years). Results are summarized in table 2.10.

Most of the exogenous parameters play a negligible role in shaping the main results. Higher (lower) risk aversion θ leads to a lower (higher) marginal utility of consumption; hence college parents, who have higher wages on average, tend to spend more (less) time with their children compared to noncollege parents when θ is higher (lower). The intergenerational correlation of cognitive skills is not affected by assumptions on the parameters, but the existence of the "Flynn Effect" would imply a higher correlation than what is observed in the NLSY/79.

2.7 Policy Experiment

In this section I explore the effects of applying the German scheme of child allowances (Kindergeld) in the US economy. The Kindergeld is one of several transfer programs, and must be distinguished from parental allowances (Elterngeld) and Maternity Pay (Mutterschaftsgeld).

The Kindergeld transfer program started in 1936; in 2012, the Kindergeld granted a monthly payment

	Baseline	$\theta = 0.8$	$\theta = 1.2$	$\sigma = 2.5$	$\sigma = 5$	Flynn Effect
Share of time α	0.77	0.64	0.88	0.77	0.76	0.74
Ratio Early/Late Time Invested	2.8	2.8	2.8	2.9	2.8	2.9
Δ College/NonCollege (early)	+7.4 %	+4.2 %	+11.0 %	+7.7 %	+7.4 %	+6.5 %
Δ College/NonCollege (late)	+12.7 %	+8.9 %	+16.8 %	+13.1 %	+12.8 %	+10.6 %
Ratio Early/Late Hours of Work	89.3 %	90.1 %	88.8 %	91.2 %	87.4 %	89.6 %
Interg. Correlation of Cog. Skills	0.45	0.45	0.45	0.45	0.45	0.50
Interg. Correlation of NonCog. Skills	0.19	0.19	0.19	0.19	0.19	0.19
Corr. Cog Skills / NonCog. Skils	0.26	0.27	0.27	0.27	0.27	0.24
Corr. NonCog Skills / HH Income	0.31	0.30	0.31	0.31	0.31	0.33

Table 2.10: Sensitivity Analysis: effect of changes in externally set parameters on Share of time, Ratio of Early to Late child care time, Difference between College and NonCollege time, Ratio of Early to Late hours of work, Intergenerational correlation of skills, final correlation between offsprings' cognitive and noncognitive skills, final correlation between offspring's skills and Income.

of 184 euros per child to virtually all households who have a child under the age of 18, although it can be extended to age 25 if the child is in school, at university or is doing professional training. The payment is performed for each child in the household, and raises to 190 euros for the third child and 215 for each additional child. The payment extends to citizens of EU countries and of several other countries, provided that they reside in Germany, and is not means-tested[21].

The 2012 Kindergeld for the first and second child amounted to approximately 5 % of the average household income in Germany[22]. For simulation purposes, I perform a flat transfer of 5 % of the model-generated average income to all households from birth of the child to age 14 and compare the effect of such policy with respect to the baseline model. Results for aggregate averages are summarized in table 2.11.

As expected, the transfer has a negative effect on labor supply, which drops by almost 4 % during early childhood and by 3 % during late childhood; however, additional income has a positive effect on

[21] Source: Social Security Throughout the World (http://www.ssa.gov/policy/docs/progdesc/ssptw/).

[22] In 2012, yearly average household income in Germany was 43500 euros. Source: http://www.voxeu.org/article/are-germans-poorer-other-europeans-principal-eurozone-differences-wealth-and-income, data from ECB Household Survey 2013.

Aggregate Effects	Baseline	Policy		
Final Cog Skills				
Mean		+ 3.0 %		
Std. Dev.		+ 1.4 %		
Final NonCog Skills				
Mean		+ 3.5 %		
Std. Dev.		+ 1.8 %		
Correlation($s_{C,T+1}, s_{C,P}$)	0.454	-0.2 %		
Correlation($s_{N,T+1}, s_{N,P}$)	0.186	-1.8 %		
Correlation($s_{C,T+1}$, HH Income)	0.306	-5.1 %		
Correlation($s_{N,T+1}$, HH Income)	0.194	-10.8 %		
	First Stage		Second Stage	
	Baseline	Policy	Baseline	Policy
Portion of Transfer Consumed		94.1 %		98.4 %
Consumption, Avg.		+ 2.9 %		+ 3.2 %
Hours Worked, Avg.	0.302	-3.7 %	0.333	-2.7 %
Time Investment, Avg	0.092	+ 4.8 %	0.033	+ 3.7 %

Table 2.11: Changes from baseline to flat transfer of 5 % of average income; aggregate variables. Changes in skills are reported as percentage of a standard deviation of baseline skills.

the time investment of households, increasing it by almost 5 % during early childhood and by 4 % in late childhood. As a result, average final cognitive achievements of children increase by 3 % of a standard deviation, while noncognitive achievements increase by 3.5 %. A closer look at how gains are distributed across households (table 2.12) reveals heterogeneity in the returns of the policy.

Stronger effects can be observed among low income households (first decile of the income distribution), who experience gains in cognitive/noncognitive skills of their offspring as large as 5 % and 7 % of a standard deviation, respectively. This is because the transfer has a larger negative effect on the labor supply of low-income households, freeing up resources and time for children; as a result, the correlation between parental income and final cognitive skills of the offspring decreases by 5.1 %, and that with noncognitive skills by almost 11 %.

The table also provides an intuition for why the intergenerational correlation of cognitive skills does not change noticeably when the policy is applied. When the transfer is performed, low-income households increase their time investment by more than higher-skilled households, because the latter face a higher opportunity cost of time. However, differences in absolute levels of investment between skills groups are smaller, because higher-skilled groups have higher wages, hence they use more resources for investment.

	D1	Q1	Q2	Q3	Q4
Cognitive Skills					
by Child's Cog Skill	3.0 %	3.0 %	3.0 %	3.1 %	3.1 %
by Child's NCog Skill	2.8 %	2.9 %	3.0 %	3.1 %	3.2 %
by Parental Cog Skill	1.8 %	2.2 %	3.0 %	3.4 %	3.6 %
by Parental NCog Skill	2.8 %	2.9 %	3.0 %	3.0 %	3.2 %
by Income	4.9 %	4.3 %	3.2 %	2.7 %	2.0 %
NonCognitive Skills					
by Child's Cog Skill	3.5 %	3.5 %	3.5 %	3.5 %	3.4 %
by Child's NCog Skill	3.2 %	3.3 %	3.4 %	3.6 %	3.7 %
by Parental Cog Skill	4.3 %	4.0 %	3.6 %	3.4 %	2.9 %
by Parental NCog Skill	3.5 %	3.4 %	3.5 %	3.5 %	3.5 %
by Income	7.0 %	5.7 %	3.7 %	2.7 %	1.8 %
1st Stage Time Invested (percentage)					
by Child's Cog Skill	5.01 %	5.01 %	5.02 %	4.96 %	4.90 %
by Child's NCog Skill	5.52 %	5.37 %	5.06 %	4.89 %	4.58 %
by Parental Cog Skill	7.67 %	6.79 %	5.23 %	4.42 %	3.45 %
by Parental NCog Skill	6.14 %	5.82 %	5.14 %	4.83 %	4.17 %
by Income	11.68 %	9.03 %	5.13 %	3.57 %	2.16 %
2nd Stage Time Invested (percentage)					
by Child's Cog Skill	3.84 %	3.83 %	3.81 %	3.75 %	3.68 %
by Child's NCog Skill	4.10 %	4.00 %	3.82 %	3.72 %	3.52 %
by Parental Cog Skill	5.60 %	5.01 %	3.97 %	3.40 %	2.68 %
by Parental NCog Skill	4.60 %	4.37 %	3.89 %	3.66 %	3.19 %
by Income	8.70 %	6.76 %	3.90 %	2.73 %	1.67 %
1st Stage Investment (absolute)					
by Parental Cog Skill	0.11	0.10	0.09	0.08	0.07
by Parental NCog Skill	0.10	0.10	0.09	0.09	0.08
by Income	0.16	0.13	0.09	0.07	0.05
2nd Stage Investment (absolute)					
by Parental Cog Skill	0.03	0.03	0.02	0.02	0.02
by Parental NCog Skill	0.02	0.02	0.02	0.02	0.02
by Income	0.04	0.04	0.03	0.02	0.01
1st Stage Hours Worked					
by Parental Cog Skill	-5.4 %	-4.9 %	-3.8 %	-3.3 %	-2.6 %
by Parental NCog Skill	-4.4 %	-4.2 %	-3.8 %	-3.5 %	-3.1 %
by Income	-8.2 %	-6.5 %	-3.8 %	-2.7 %	-1.7 %
2nd Stage Hours Worked					
by Parental Cog Skill	-4.0 %	-3.6 %	-2.8 %	-2.4 %	-1.9 %
by Parental NCog Skill	-3.2 %	-3.1 %	-2.7 %	-2.6 %	-2.3 %
by Income	-6.1 %	-4.8 %	-2.8 %	-1.9 %	-1.2 %

Table 2.12: Changes from baseline to flat transfer of 5 % of average income; changes in final cognitive and noncognitive skills as percentage of the baseline standard deviation.

Finally, higher-skilled parents are still more productive than low-skilled parents. As a result, cognitive skills of children of high-skilled households increase by more compared to the children of low-skilled households, so that the policy has little effect on the intergenerational persistence of cognitive skills.

This result may help understand why in Germany the intergenerational correlation of test scores is similar to that of the US (Anger and Heineck (2009)) although the US have no child allowance policy. Child allowances can reduce the impact of income by itself, that is the fact that higher-income households have more resources for their children, but cannot influence the fact that higher-skilled parents are more productive in raising skillful children.

2.8 Conclusions

I develop a model of skills formation and household choices, grounded in the literature on Cognitive and NonCognitive Skills, and I show that it can help explaining several stylized facts on child care time and cognitive/noncognitive achievement. Results suggest that differences in returns to investment between early childhood and late childhood can explain why child care time of children aged 0-5 is more than twice as large as later child care time and a sizeable fraction of the differential in child care time between college-educated and noncollege-educated parents. Finally, the model is also capable of replicating endogenously the observed pattern of correlations between parental skills and cognitive/noncognitive achievements of their children, and accounts for a sizeable portion of the intergenerational correlation of income.

Results are robust to a number of alternative assumptions on externally set parameters; I also show that a similar cognitive-skills-only model is not able to capture several important features of the data.

I simulate the effect of applying the 2012 German scheme of child allowances in the US economy; results show that such policy could improve average cognitive skills by 3 % and noncognitive skills by 3.5 %, and reduce the correlation between cognitive skills of a child and household income by 5 %. However, the policy has a negligible effect on the intergenerational persistence of cognitive skills; such result is compatible with the fact that the mother-offspring correlation of cognitive achievement is of similar magnitude in the US and in Germany.

Appendix

2.A Analytical Results

First order conditions for consumption c_t give

$$c_t^{-\theta} = \lambda_t$$

where λ_t is the multiplier associated to the budget constraint of the household. First order conditions with respect to labor time n_t gives

$$(1 - n_t - x_t) = \left(\frac{\zeta}{w\lambda_t}\right)^{1/\sigma}$$

substituting the first equation inside the second yields that households trade off leisure and consumption according to the equation

$$(1 - n_t - x_t) = \left(\frac{\zeta c_t^{\theta}}{w}\right)^{1/\sigma}$$

Taking first order conditions with respect to e_t yields

$$\lambda_t = \mu_t A(1 - \alpha)(\frac{x_t}{e_t})^{\alpha}$$

substituting the multiplier λ_t and dividing the last expression by 2.7 yields

$$e_t = w(\frac{1 - \alpha}{\alpha})x_t \tag{2.13}$$

2.B Relations between investment and parameters

Consider the multipliers μ_t associated to the constraint $I_t = Ax_t^{\alpha}e_t^{1-\alpha}$ as the function

$$\mu_t = \mu_t(K_t, S_t, x_t)$$

where the function expresses the marginal productivity of investment as a function of the stage t, of the parameters K_t encompassing the efficiency of investment[23], of the state S_t and of the amount of time spent with the child x_t. By definition $\partial\mu_t/\partial K_t > 0$; moreover, since the technology has decreasing returns to every single input, $\partial\mu_t/\partial x_t < 0$.

Proposition 1 (1)**.** *Suppose that*

- ***A1****:* $\alpha > 1/2$*;*

then we have that $\frac{\partial x_t}{\partial K_t} > 0$*, that is, households respond to increased productivity in investment by increasing time invested in the offspring.*

Furthermore, if we have that

- ***A2****: Preferences satisfy Balanced Growth Path, that is* $\theta = 1$*;*
- ***A3****:* $\sigma \in [0, 1], \zeta > \frac{1-\alpha}{\alpha\sigma}$*;*

then $\frac{\partial n_t}{\partial K_t} < 0$*, that is households respond to higher productivity in investment by decreasing hours of work.*

Proof. First of all, by implicit function theorem we have

$$\frac{\partial F}{\partial n_t}\frac{\partial n_t}{\partial K_t} + \frac{\partial F}{\partial x_t}\frac{\partial x_t}{\partial K_t} = -\frac{\partial F}{\partial K_t}$$

Call equation 2.5 F_1 and equation 2.6 F_2; substituting expression 2.8 for e_t and the budget constraint in equation 2.5 and applying the result above yields

[23] In practice, many coefficients may enter K_t; for instance, in a Cobb-Douglas specification of the technology of skills formation, K_t includes the scale of the function and the exponent of the investment variable. In a CES specification, K_t includes the coefficient that multiplies the investment variable inside the CES aggregator.

$$\left[-\sigma\zeta(1-n_t-x_t)^{\sigma-1}-\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}\right]\frac{\partial n_t}{\partial K_t}+$$
$$\left[-\sigma\zeta(1-n_t-x_t)^{\sigma-1}+\frac{1-\alpha}{\alpha}\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}\right]\frac{\partial x_t}{\partial K_t}=0$$
$$\left[\sigma\zeta(1-n_t-x_t)^{-\sigma-1}\right]\frac{\partial n_t}{\partial K_t}+\left[\sigma\zeta(1-n_t-x_t)^{-\sigma-1}\right]\frac{\partial x_t}{\partial K_t}=(\frac{\partial\mu_t}{\partial K_t}-\frac{\partial\mu_t}{\partial x_t})A\alpha^{\alpha}w^{1-\alpha}(1-\alpha)^{1-\alpha}$$

Trivial manipulation of the two equations yields

$$\frac{\partial n_t}{\partial K_t}=-\frac{\partial x_t}{\partial K_t}\frac{\sigma\zeta(1-n_t-x_t)^{\sigma-1}-\frac{1-\alpha}{\alpha}\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}{\sigma\zeta(1-n_t-x_t)^{\sigma-1}+\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}$$
$$\frac{\partial n_t}{\partial K_t}=-\frac{\partial x_t}{\partial K_t}+(\frac{\partial\mu_t}{\partial K_t}-\frac{\partial\mu_t}{\partial x_t})\frac{A\alpha^{\alpha}w^{1-\alpha}(1-\alpha)^{1-\alpha}(1-n_t-x_t)^{\sigma+1}}{\sigma\zeta}$$

Solving the system for $\frac{\partial x_t}{\partial K_t}$ we have

$$\frac{\partial x_t}{\partial K_t}\left[1-\frac{\sigma\zeta(1-n_t-x_t)^{\sigma-1}-\frac{1-\alpha}{\alpha}\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}{\sigma\zeta(1-n_t-x_t)^{\sigma-1}+\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}\right]=$$
$$(\frac{\partial\mu_t}{\partial K_t}-\frac{\partial\mu_t}{\partial x_t})\frac{A\alpha^{\alpha}w^{1-\alpha}(1-\alpha)^{1-\alpha}(1-n_t-x_t)^{\sigma+1}}{\sigma\zeta}$$

It is clear that the right hand side is always positive, being the product of only positive terms, because $\frac{\partial\mu_t}{\partial K_t}>0$ by construction and $\frac{\partial\mu_t}{\partial x_t}<0$ because of the CES form of the skills formation technology and of the utility of skills. Hence, we have that $\frac{\partial x_t}{\partial K_t}>0$ if and only if

$$\frac{\sigma\zeta(1-n_t-x_t)^{\sigma-1}-\frac{1-\alpha}{\alpha}\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}{\sigma\zeta(1-n_t-x_t)^{\sigma-1}+\theta(wn_t-w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}<1$$

and since leisure $(1-n_t-x_t)$ is always positive and consumption $wn_t-w(\frac{1-\alpha}{\alpha})x_t$ is always bigger than zero (by the fact that marginal utility approaching zero is infinity), this can be trivially shown to be equivalent to

$$\alpha>1/2$$

which proves the first statement. Suppose now that $\alpha>1/2$, that is $\frac{\partial x_t}{\partial K_t}>0$; by the previous equation,

$$\frac{\partial n_t}{\partial K_t} = -\frac{\partial x_t}{\partial K_t}\left[\frac{\sigma\zeta(1-n_t-x_t)^{\sigma-1} - \frac{1-\alpha}{\alpha}\theta(wn_t - w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}{\sigma\zeta(1-n_t-x_t)^{\sigma-1} + \theta(wn_t - w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1}}\right]$$

Since the denominator is positive, $\frac{\partial n_t}{\partial \mu_t} < 0$ if and only if

$$\sigma\zeta(1-n_t-x_t)^{\sigma-1} > \frac{1-\alpha}{\alpha}\theta(wn_t - w(\frac{1-\alpha}{\alpha})x_t)^{\theta-1} \tag{2.14}$$

and in the case of Balanced Growth path ($\theta = 1$), a sufficient condition for 2.14 to hold is that $\sigma \in [0,1], \zeta > \frac{1-\alpha}{\alpha\sigma}$, completing the proof. □

It is not possible to state the proposition generally for any value of the concavity of leisure and any risk aversion parameter; however, condition 2.14 suggests that if the share of goods $1-\alpha$ in the investment function is low and ζ is higher than 1, the solution should satisfy the condition in the majority of cases for reasonable values of labor supply and time spent with children, but may possibly be violated in extreme cases.

2.C The Technology of Skills Formation - Parameters

Covariance Matrix				
	Child's Cog. Skills at birth	Child's NonCog. Skills at Birth	Mother's Cog. Skills	Mother's Noncog. Skills
Child Cog. Skills	0.1777			
Child NonCog. Skills	-0.0204	0.2002		
Mother's Cognitive	0.0182	0.0592	0.5781	
Mother's Noncognitive	0.0050	0.0261	0.0862	0.0667

Correlation Matrix				
	Child's Cog. Skills at birth	Child's NonCog. Skills at Birth	Mother's Cog. Skills	Mother's Noncog. Skills
Child Cog. Skills	1.0000			
Child NonCog. Skills	-0.1081	1.0000		
Mother's Cognitive	0.0569	0.1741	1.0000	
Mother's Noncognitive	0.0463	0.2260	0.4390	1.0000

Table 2.13: Variance/Covariance Matrix and Correlation Matrix for initial conditions of parental and offspring's skills [source: Cunha, Heckman, Schennach (2010), Appendix].

Technology of **Cognitive** Skills Formation				
		1st Stage		2nd Stage
Cog Skills	$\gamma_{1,C,1}$	0.485	$\gamma_{2,C,1}$	0.884
NonCog Skills	$\gamma_{1,C,2}$	0.062	$\gamma_{2,C,2}$	0.011
Investment	$\gamma_{1,C,3}$	0.261	$\gamma_{2,C,3}$	0.044
Parental Cog	$\gamma_{1,C,4}$	0.035	$\gamma_{2,C,4}$	0.051
Parental NonCog	$\gamma_{1,C,5}$	0.157	$\gamma_{2,C,5}$	0.011
Complementarity	$\phi_{1,C}$	0.585	$\phi_{2,C}$	-1.220
Elasticity of Substitution $1/(1-\phi)$		2.409		0.450
Variance of Shocks	$\eta_{1,C}$	0.165	$\eta_{2,C}$	0.098
Technology of **NonCognitive** Skills Formation				
		1st Stage		2nd Stage
Cog Skills	$\gamma_{1,N,1}$	0.000	$\gamma_{2,N,1}$	0.002
NonCog Skills	$\gamma_{1,N,2}$	0.602	$\gamma_{2,N,2}$	0.857
Investment	$\gamma_{1,N,3}$	0.209	$\gamma_{2,N,3}$	0.104
Parental Cog	$\gamma_{1,N,4}$	0.014	$\gamma_{2,N,4}$	0.000
Parental NonCog	$\gamma_{1,N,5}$	0.175	$\gamma_{2,N,5}$	0.037
Complementarity	$\phi_{1,N}$	-0.464	$\phi_{2,N}$	-0.522
Elasticity of Substitution $1/(1-\phi)$		0.683		0.657
Variance of Shocks	$\eta_{1,N}$	0.203	$\eta_{2,N}$	0.102

Table 2.14: The Technology for Cognitive and Noncognitive Skill Formation; estimated by Cunha, Heckman, Schennach (2010) [pag. 919] taking into account Investment endogeneity; skills linearly anchored to educational attainment; factors normally distributed, standard errors in parentheses.

Technology of **Cognitive** Skills Formation				
		1st Stage		2nd Stage
Cog Skills	$\gamma_{1,C,1}$	0.303	$\gamma_{2,C,1}$	0.448
NonCog Skills	$\gamma_{1,C,2}$	0	$\gamma_{2,C,2}$	0
Investment	$\gamma_{1,C,3}$	0.319	$\gamma_{2,C,3}$	0.098
Parental Cog	$\gamma_{1,C,4}$	0.378	$\gamma_{2,C,4}$	0.454
Parental NonCog	$\gamma_{1,C,5}$	0	$\gamma_{2,C,5}$	0
Complementarity	$\phi_{1,C}$	-0.180	$\phi_{2,C}$	-0.781
Elasticity of Substitution $1/(1-\phi)$		0.847		0.561
Variance of Shocks	$\eta_{1,C}$	0.193	$\eta_{2,C}$	0.050

Table 2.15: The Technology for Only Cognitive Skill Formation (used for counterfactual experiment); estimated by Cunha, Heckman, Schennach (2010) [Online Appendix] taking into account Investment endogeneity; skills linearly anchored to educational attainment; factors normally distributed, standard errors in parentheses.

2.D Solution Algorithm

The algorithm uses first order conditions when possible and solves the household's problem backwards, using polynomial approximation of next period's value function. Shocks are approximated with a three-state symmetric shock, that has mean zero and variance as in table 2.14. The reason is that shocks are independent, hence for n-state shocks the next period's value function must be computed n^2 times, greatly increasing the computational burden. A brief description of the algorithm follows:

1. Start from final period $t = T + 1$; extract 10^5 points in the continuous state space of $S_{T+1} = \{s_{C,T+1}, s_{N,T+1}, s_{C,P}, s_{N,P}, \epsilon_P^w\}$, distributed as uniform Sobol numbers from -10 to +10 standard deviations of the covariance matrix in table 2.13. The algorithm oversamples the tails on purpose (the true distribution is jointly normal at the initial period, and roughly normal in later periods), because it is more difficult to approximate the value function far from the median. Also, simulations show that the variance of offspring's skills increases by approximately 70 % from period 1 to period $T+1$, so it is important to consider a more dispersed final distribution. I use Sobol numbers because they span the state space more efficiently than uniform random numbers.

2. Given the state, the solution at period t is found as follows. The objective function at time t, before maximization, is

$$Q_t(S_t, c_t, n_t, e_t, x_t) = \frac{c_t^{1-\theta}}{1-\theta} + \zeta \frac{(1-n_t-x_t)^{1-\sigma}}{1-\sigma} + W(s_{C,t}, s_{N,t}) + \beta \mathbb{E}\Big[V_{t+1}(S_{t+1})\Big]$$

which is a function of the state $S_t = \{s_{C,t}, s_{N,t}, s_{C,P}, s_{N,P}, \epsilon_P^w\}$ and of the controls c_t, n_t, x_t, e_t. Clearly, $V_t = \max_{\{c_t,n_t,x_t,e_t\}} Q_t(S_t, c_t, n_t, x_t, e_t)$, subject to all the constraints of the dynamic problem in section 3. In all periods, the choices x_t and e_t map offspring's skills $s_{C,t}, s_{N,t}$ into next period's skills $s_{C,t+1}, s_{N,t+1}$, hence in the state S_{t+1} and in the value function $V_{t+1}(S_{t+1})$. Then it is necessary to predict next period's value function. At period $T+1$ such prediction is unnecessary, because $V_{T+2} = 0$. Assume that the map from the state S_{t+1} to the value function $V_{t+1}(S_{t+1})$ is known.

3. Start with a guess x_t. Given x_t and the state S_t, one of the controls e_t is the explicit solution to equation 2.7 in the optimum; call it $e_t^*(x_t)$. Given $e_t^*(x_t)$, consumption c_t can be backed out from the budget constraint given the labor choice n_t. Then n_t is the solution to the equation 2.5, which is solved by bisection method; denote these last two choices as $c_t^{**}(x_t)$,$n_t^{**}(x_t)$. As a result, given the state S_t, now Q_t is only a function of the state S_t and of the choice of x_t, because all other controls are functions of x_t. Denote this objective function as $Q_t(S_t, c_t^{**}(x_t), n_t^{**}(x_t), e_t^*(x_t), x_t)$. Notice that $Q_t(S_t, c_t^{**}(x_t), n_t^{**}(x_t), e_t^*(x_t), x_t)$ is not equal to $\max_{c_t,n_t,e_t} Q_t(S_t, c_t, n_t, e_t, x_t)$ for all x_t, except in the optimum x_t^*, because $c_t^{**}, n_t^{**}, e_t^*$ are not the optimal controls given a suboptimal x_t since they exploit the first order conditions of the problem, which are jointly true only in the optimum. However,

$$Q_t(S_t, c_t^{**}(x_t), n_t^{**}(x_t), e_t^*(x_t), x_t) \leq \max_{c_t,n_t,e_t} Q_t(S_t, c_t, n_t, e_t, x_t)$$

because, for given x_t, all other controls are suboptimal. Finally,

$$\max_{x_t} Q_t(S_t, c_t^{**}(x_t), n_t^{**}(x_t), e_t^*(x_t), x_t) = \max_{c_t,n_t,e_t,x_t} Q_t(S_t, c_t, n_t, e_t, x_t)$$

which ensures that the solution is the same. Then, $\tilde{V}_t(x_t, S_t)$ is maximized on x_t using golden search. The algorithm gives the solution at arbitrary precision for every state point; the chosen precision is 10^{-7}. One possible concern is that, given the transformation described above, the function might lose single-peakedness, which is a strict requirement of golden search. Using another algorithm, i.e. grid-based maximization methods, delivers substantially the same results but is much slower and produces greater approximation errors. Finally,

$$V_t(S_t) = \max_{c_t, n_t, e_t, x_t} Q_t(S_t, c_t, n_t, e_t, x_t)$$

so the maximizers for Q_t given S_t deliver the desired policy functions.

4. Now we have 10^5 values of the value function $V(s_{C,t}, s_{N,t}, s_{C,P}, s_{N,P}, \epsilon_P^w)$ associated to state points S_t. Name X_t the $10^5 \times K$ matrix which stores an n-th degree polynomial[24] in the values of the extracted state points at time t. To explain further, X_t includes columns of 10^5 realizations of $1, s_{C,t}, s_{C,t}^2, s_{C,P}, s_{C,P}^2, s_{C,t}s_{C,P}, ...$ and so on. A polynomial regression is performed:

$$V_t = X_t \beta_{V_t} + \epsilon_t \tag{2.15}$$

which coefficients are stored in the $K \times 1$ vector β_{V_t}.

5. Now go back one period and back to point 2, predicting next-period value function V_{t+1} by using $\beta_{V_{t+1}}$. The algorithm stops at the solution of period 2 because predicting V_1 is unnecessary.

After the household's problem is solved, I extract random jointly lognormal initial conditions for every household using the variance/covariance matrix 2.13 and I solve again each household's problem, starting in period 1, to obtain policy functions. The reason is that the value function is easier to approximated by polynomial approximation than the policy functions for n_t, x_t, e_t.

The most nonstandard part of the algorithm is the polynomial approximation, which has to be reliable in order not to produce large errors in the solution. In this case the approximation is quite precise; define the relative ex-post approximation error ϵ^{approx} as

$$\epsilon^{\text{approx}} = \frac{\hat{V}_t(S_t) - V_t(S_t)}{\text{std}(V_t(S_t))}$$

that is, the difference between the ex-ante prediction of the value function $\hat{V}_t(S_t)$ and the ex-post solution to the problem $V_t(S_t)$, found with maximization, normalized by the standard deviation of $V_t(S_t)$.

[24]It follows that the number of terms K is given by $\binom{n+5-1}{n}$, where n is the degree of the polynomial and 5 is the number of variables.

The average of such relative error is always under 10^{-4} for every t and the maximum relative error is under 10^{-2}.

2.E Correlation Matrices of Data Factors

	Cog. Skills at age 6	NonCog. Skills at age 6	Mom's Cog. Skills	Mom's NonCog. Skills
Cog. Skills at age 6	1.0000			
NonCog. Skills at age 6	**0.1566***	1.0000		
	0.0007			
Mom's Cog. Skills	**0.2836***	0.2097*	1.0000	
	0.0000	0.0000		
Mom's NonCog. Skills	0.1528*	0.2021*	0.4167*	1.0000
	0.0000	0.0000	0.0000	

Table 2.16: Pairwise Correlation Matrix of Skills Factors, with significance level (* = 0.001 significance); all families (minimum N = 460). Calibration targets are displayed in bold. Source: data extract from NLSY79/CNSLY79 by Cunha, Heckman and Schennach (2010).

	Cog. Skills at age 14	NonCog. Skills at age 14	Mom's Cog. Skills	Mom's NonCog. Skills
Cog. Skills at age 14	1.0000			
NonCog. Skills at age 14	0.2494*	1.0000		
	0.0000			
Mom's Cog. Skills	0.4303*	0.1336*	1.0000	
	0.0000	0.0000		
Mom's NonCog. Skills	0.1844*	0.1444*	0.4167*	1.0000
	0.0000	0.0000	0.0000	

Table 2.17: Pairwise Correlation Matrix of Skills Factors, with significance level (* = 0.001 significance); all families (minimum N = 1009). Source: data extract from NLSY79/CNSLY79 by Cunha, Heckman and Schennach (2010).

2.F Data on Child Care Time

The data of Ramey and Ramey (2010) consist in 13 time diary surveys for US, for years 1965, 1975, 1985, 1992-1994, 1995, 1998, 2000, and all years from 2003 to 2008. I use the time diaries for 1975 and 1985 to compute the averages with which I compare the model. The online Appendix of Ramey and Ramey provides details on how child care time is computed, i.e. which activity codes have been included in total child care time.

I compute the averages selecting only married parents; hence I use a total of 2837 observations. Of these, 1312 (46 %) are males. College-educated individuals are 624 (22 %), of which 355 are males (57 % of college-educated). Of all parents, 1218 (43 %) have a child under the age of 5. I use the AHTUS recommended weights for computing averages. Results for averages and ratios are robust to the inclusion of older individuals.

I calculate two limiting cases: the "college-mating" case in which college mothers are always married to college fathers, and the "random-mating" case in which college mothers are randomly mating with all fathers. Denote the averages for different cases as $\bar{x}_{i,j,k}$ where $i \in \{1,2\}$ stands for the stage of development of the child (1 for ages 0-5, 2 for subsequent ages), $j \in \{f,m\}$ stands for the sex of the parent and $k \in \{C, NC, ALL\}$ stands for college/noncollege/all. Then the household-level averages for stage i and education k are obtained by

$$\begin{aligned}\bar{x}_{i,k}^{\textbf{college-mating}} &= \bar{x}_{i,f,k} + \bar{x}_{i,m,k} \\ \bar{x}_{i,k}^{\textbf{random-mating}} &= \bar{x}_{i,f,k} + \bar{x}_{i,m,ALL}\end{aligned}$$

	Baseline	Final Preferences	Equal Elasticity = 0	Two Periods
Share of time α	0.77	0.67	0.76	0.76
Ratio Early/Late Time Invested	2.8	0.9	6.4	3.2
Δ College/NonCollege (early)	+7.4 %	+4.7 %	+8.1 %	+15.9 %
Δ College/NonCollege (late)	+12.7 %	+ 3.5 %	+13.9 %	+14.0 %
Ratio Early/Late Hours of Work	89.3 %	104.8 %	87.4 %	90.6 %
Interg. Correlation of Cog. Skills	0.45	0.42	0.36	0.33
Interg. Correlation of NonCog. Skills	0.19	0.17	0.19	0.19
Corr. Cog Skills / NonCog. Skils	0.26	0.17	0.22	0.21
Corr. NonCog Skills / HH Income	0.31	0.28	0.26	0.25

Table 2.18: Robustness Checks: effect of changes in model structure and technology assumptions on Share of time, Ratio of Early to Late child care time, Difference between College and NonCollege time, Ratio of Early to Late hours of work, Intergenerational correlation of skills, final correlation between offsprings' cognitive and noncognitive skills, final correlation between offspring's skills and Income.

Chapter 3

Parental Links and Employment Prospects: Evidence from the UK

3.1 Introduction

We explore how parental links affect labor market choices and employment prospects of individuals; more specifically, we look at how the job finding and job separation rate are affected by the fact that parents are employed rather than unemployed, and we investigate patterns of occupational mobility across generations. We construct detailed monthly job histories using information from the British Household Panel Survey and exploit information on household structure and friends from the same dataset. In particular, we improve upon the existing literature by linking each individual to his family members, and measuring how the employment prospects of an individual are affected by the labor market status of his relatives or his spouse. Finally, we compare the strength of family ties with that of other relevant ties such as friends.

Our contribution is twofold: first, we document the extent of intergenerational occupational mobility and we argue that using contemporaneous information on parental occupation, rather than retrospective information, is important for measuring such mobility. Second, we establish that the father's employment has an impact on the offspring job finding probability, and we argue that such impact is due to parental networks being an important determinant of employment prospects.

The importance of social networks in determining labor market outcomes has been recognized in the literature in the last decades[1]. Networks are a common way to alleviate information frictions, largely

[1] Rees (1966) and Granovetter (1973) were the first ones to investigate the important role played by social networks

used by both workers[2] and firms[3].

Our work looks at the workers' side, i.e. individual transitions from unemployment to employment and viceversa. Several papers have tried to quantitatively assess how belonging to a particular network affects labor market variables such as job finding, job separation and wages[4]. One usual shortcoming of the data is that only indirect measures of networks are available. Some researchers rely on estimates of social networks, in order to assess their impact. As a consequence, the estimates produced by these studies are likely to be affected by measurement error, due to the impossibility of exactly identifying the network members. Our work studies social networks at a disaggregated scale and uses direct information on social contacts, thus limiting the extent of measurement error.

The main focus of this work is on parental links. The exogenous nature of parental links –individuals do not choose their parents– makes it easier to quantify their effects and reduces problems of double causality. Fathers and mothers are also commonly recognized to be strong ties in the network literature, and it is therefore interesting to analyze the extent to which parents can influence the labor market choices and outcomes of their offspring.

Moreover, such influence is likely to affect the intergenerational persistence of social and economic status. In this sense, the choice of the data is particularly appealing for our analysis: among developed countries, the UK ranks relatively high in terms of socio-economic persistence across generations[5].

In order to motivate our econometric specification, we first postulate a stylized model of intergenerational transmission of networks in the labor market. In the model, offspring inherit a fraction of the father's network, and then develop their own contacts while employed. As they spend more and more time in employment, the correlation between their employment status and the father's one fades out over time. This motivates our empirical strategy based on difference-in-differences estimation, employing a threshold age to distinguish between treatment and control group. We also report the results of other linear probability models, controlling for individual fixed effects. We show that the effect of parental links is larger when the individual looks for a job in the same occupation of the father. Although one might think that parental links play the most important role in helping the offspring to find his very first

in labor markets. Montgomery (1991) proposed a simple model to capture the features of a labor market with personal connections. More recently, Calvó-Armengol & Jackson (2004) studied the dynamic implications of networks, shedding new light on the possible effects of policy.

[2]See for instance Holzer (1988) and Pellizzari (2010).

[3]See for instance Ioannides & Loury (2004) and Topa (2011).

[4]See for instance Topa (2001), Munshi (2003), Beaman (2012).

[5]The intergenerational earnings elasticity in UK is estimated to be about 0.5, one of the highest among developed countries, again very similar to that of the US, which ranges from 0.5 to 0.6 depending on the estimation method (Corak (2006)).

job, we document that large and persistent differences in the job finding are related to father's labor market variables for a number of years, rather than only at the beginning of one's career.

To the best of our knowledge, we are the first ones to analyze how parental links affect transitions in and out of unemployment, rather than non-employment as other studies in the literature. We are also the first to document the existence of a strong positive effect of father's employment on these transitions by exploiting direct information on such links. We choose to look at employment and unemployment because the choice of participating to the labor force can be influenced by parental background and employment, which in turn would confound the effect of parental links on nonemployment to employment transitions. By looking at unemployment vs. employment, we are selecting those individuals who are participating to the labor force to begin with.

We document that in the UK occupations tend to be persistent across generations; for instance, sons are from 26 % to 167 % (depending on the sector) more likely to end up working in similar occupations as their fathers, with some exceptions. Similar considerations apply to daughters and mothers.

We find that having an employed (rather than unemployed) father increases the employment rate by about 8 percentage points, with an effect on the job finding rate of at least 5 percentage points, compared to an average in-sample job finding rate of 11 %. Moreover, if an individual searches for a job in the same sector in which his father is currently employed, the effect on the job finding rate is magnified by a further 4 percentage points. Such results are robust to alternative specifications and to several robustness checks. Overall, the evidence is strongly consistent with our model of intergenerational networks. Moreover, by means of a number of empirical tests we are able to rule out several other possible mechanisms that could potentially generate our results. We do not find similar effects for mothers. For the sake of comparison, an additional employed friend increases the job finding rate by 1 - 3 % depending on the estimation method, while the spouse's employment status has a strong association with the individual job finding rate (with a similar magnitude to the father's one).

The rest of the paper is organized as follows. Section 3.2 surveys the literature in greater detail, emphasizing the differences between our work and the others. In Section 3.3 we introduce the data, along with some descriptive statistics of interest. In Section 3.4 we present a stylized model of intergenerational networks, in order to justify the empirical models employed for the analysis (explained in Section 3.5). Results are shown in Section 6 and discussed in Section 7. We performs some robustness checks in Section 8 and conclude in Section 9.

3.2 Related Literature

Our paper relates to the extensive literature on intergenerational occupational and income mobility. In particular, our work suggests two possible sources of income persistence across generations. One is through higher chances of being employed, and the other one is through occupational persistence. As long as wages differ across occupations, then the influence of parental background on occupational choices can be potentially very important in explaining income persistence. The literature on the persistence of income across generations dates back to Becker & Tomes (1986). Solon (1992) is one of the earliest assessments of the measurement error issues affecting the estimation of intergenerational elasticities, finding high values of persistence for the US. A comprehensive survey of the literature is provided in Corak (2006), who performs a cross-country study. Jäntti *et al.* (2006) also perform a study of the intergenerational earnings mobility across several countries, while Björklund *et al.* (2012) focuses on Sweden, with a particular emphasis on the top of the distribution. On the link between occupational and income persistence across generations, see also Corak & Piraino (2011).

We offer direct evidence of the positive impact of family ties' employment on labor market transitions. Many empirical studies try to identify the effect of belonging to a particular network on labor market outcomes. Several papers rely on indirect measures of networks. For instance, Topa (2001), Bayer *et al.* (2008) and Schmutte (2010) use geographical proximity and group affiliation as proxies for social interactions. Beaman (2012) uses data on political refugees resettled in the US and proxies for networks using nationality. Overall, these studies find evidence of positive effects of social interactions on labor market outcomes. Similarly, Khan & Lehrer (2013) use a random assignment to a unique intervention to identify the impact of changes in the size of a social network. Access to the program successfully led to gains in the number of weak ties but these changes did not translate into improved employment outcomes. Herault & Kalb (2009) look instead at parental links; using retrospective parental information from Australian data, they find significant persistence in employment across generations.

Our paper is more closely related to that strand of this literature that exploits direct identification of network members.

O'Regan & Quigley (1993) study the correlation of employment status of urban youth with the employment status of their family members (parents and siblings) in the US, finding strong and positive correlations. Further, they observe that the industry affiliation of the network members is a good predictor of the industry affiliation of the individual. Magruder (2010) examines to which extent parents help children in finding jobs in South Africa. He finds that fathers help sons (but not daughters), while mothers are not helpful in finding jobs. Differently from these works, our analysis is dynamic and focuses on transitions from unemployment to employment and viceversa, rather than on employment status versus nonemployment. Also, our data allows us to employ different estimation techniques and to compare parental effects to similar effects by other strong ties. Kramarz & Skans (2014) investigate parental networks at the firm-level. They analyze Swedish graduates, finding that it is quite frequent that their first job is in the same plant where their parents work. With respect to their paper, rather than focusing only on the entry in the labor market, we follow individuals over their life-cycle, investigating whether the advantages derived from their network persist over time.

Finally, Pistaferri (1999) uses Italian data and finds that informal networks use is associated with higher job finding rates and lower wages. Similarly, Bentolila *et al.* (2010) find that individuals who use social contacts to find their job are characterized by higher job finding rate (lower unemployment duration) and slightly lower wages. They suggest that the trade-off between job finding rate and wage could still make individuals choose to enter the same sector as their network members. Indeed, we document patterns of intergenerational persistence in occupations; along with our regression results, this is consistent with a model of occupational choice in the spirit of Bentolila *et al.* (2010).

Closely related to our work is the study of Cappellari & Tatsiramos (2011), who also use data from the BHPS and a similar methodology. Nonetheless, some relevant traits differentiate the two works: first, we focus on parental links, instead of friendship ones; second, we identify monthly transitions (rather than yearly); third, we look at transitions within the labor force while they consider transitions from non-employment to employment; fourth, we document how searching for a job in the same occupational sector magnifies the effects we find.

Several studies in the literature have tackled the issue of understanding the effects of social networks by means of a theoretical model. One of the first papers to include personal contacts in a job search framework was Mortensen & Vishwanath (1994). In their model the information about vacancies comes from two different sources: direct application to employers or indirect contact through friends. As a consequence, better connected individuals have more chances to find a job. Similarly, Montgomery (1991) finds that well connected workers perform better in the labor market, both in terms of wages and of higher job finding rates. Calvó-Armengol & Jackson (2004) also develop a model where workers can obtain information through an explicitly modeled network of social contacts. In their model, belonging to a network with less employed members implies worse employment prospects, and this effect is persistent over time. Other models of networks and job search are in Fontaine (2008) and Calvó-Armengol & Zenou (2005), with a particular focus on networks' dynamics. More recently, Galenianos (2014) embeds networks explicitly into a search and matching model and finds that referral mechanisms have important macroeconomic implications.

A distinctive feature of all these works is that networks exhibit a positive effect on labor market outcomes of individuals. These studies constitute the theoretical ground on which we base the interpretation of our results.

3.3 The Data

We use data from the British Household Panel Survey, a representative sample from the UK following individuals since 1991. The BHPS is a yearly survey taken by about 10,000 individuals per year and the last available wave for this study is 2008. The follow-up rate is very high and the great majority (more than 90%) of individuals are interviewed also in the subsequent year. Besides these, every year a certain number of new individuals enter the sample. A total of 32,377 individuals are interviewed in the BHPS in the period 1991-2008. Even though the survey is yearly, individuals report their job history in the last year, listing all the employment (unemployment) spells along with several characteristics of each job. This allows us to identify monthly transitions and build long time series for each individual, up to 216 months. Details on how we construct job histories for individuals are included in Appendix A.

We retrieve the employment status of individuals exploiting the job histories, distinguishing between employees and self employed. The employment status of individuals is assigned at the monthly frequency. Differently from other studies, we do not consider individuals who are out of the labor force in our transitions. We define the job finding rate as the probability of transiting from unemployment (rather than non-employment, as for instance in Cappellari & Tatsiramos (2011)) to employment. The job separation rate is defined accordingly. We restrict our sample to individuals aged between 16 and 65[6] and, as it is standard in the literature, we drop armed forces and registered disabled. Eventually we are left with 27,278 individuals, for a total of 2,232,528 monthly observations.[7]

[6] That is, our intergenerational sample will include couples of parents and offspring if and only if both are in this age range.

[7] We also check whether our final sample is representative of the UK economy between 1991 and 2008. We compute the in-sample unemployment rate and compare it with the harmonized unemployment rate according to OECD statistics

Along with a detailed job history, for each individual a large amount of information is typically available, including sex, age, education, occupation, race, marital status, region of residence and much more. More interestingly, the identification number of parents and spouse is available, allowing us to connect individuals to their family members and follow them together over time[8]. In addition to this, the data include information on the employment status of the three closest friends and the occupation of the closest friend. This information is collected only every second wave, starting in 1992. At the core of the analysis, we consider the relationship between the employment status (and the occupation) of the parents and the employment prospects of respondents. We also compare parental effects to similar effects by spouse and friends. Since friends' job histories are not reported, in order to keep the monthly frequency we extend their employment status and occupation in the following 12 months after each observation. Furthermore, we use a simple procedure to attribute the occupational sector to unemployed and to extend non-varying or spell-dependent variables, as described in Appendix A. Especially for the occupation, the data contain many missing values, both for respondents and connected individuals: we assume that these values are missing at random and simply exclude the incomplete observations from the estimation, when it is not possible to replace them according to the procedure described in Appendix A. The final size of the estimating sample varies, depending on the dependent variable we use in each regression[9].

Table 3.14 (in Appendix 3.A.5) displays some descriptive statistics of interest for our analysis. As we can see, the period 1991-2008 is characterized by a relatively low level of unemployment in the UK. More than 93% of the population in the labor market has a job, 12% of which is self employed. The average monthly job finding rate, that is the probability of transition from unemployment to employment, is slightly above 7%. Conversely, the average job separation rate is relatively small (0.4%). This implies that in the period considered the UK economy was characterized by a high level of security for those who had a job. On the other hand, it was somewhat hard to find an occupation for those who were unemployed: on average, the expected waiting time in unemployment was about one year. In other words, the reason behind the low unemployment rate is the tightness of the monthly outflow from employment rather than a large inflow from unemployment. Compared to other OECD countries, the UK economy has an average performance in terms of search variables[10]. A comparison among genders

(figure 3.6 in the Appendix). The average unemployment rate of our sample replicates quite well the pattern of the OECD series.

[8]Unfortunately, it is possible to do so only for about 19% of the whole sample for fathers and 25% for mothers (those who report the parental PID numbers).

[9]By construction, the job finding rate (job separation) is defined only over unemployment (employment) spells.

[10]For a cross-country comparison of estimates of the standard search variables see Hobijn and Sahin (2007). As shown

shows that females in the labor force tend to have slightly better outcomes than males. About 53% of our sample is female and the average individual is aged 39. We define four educational groups and nine occupational sectors, following the SOC aggregation by major group, as established by the Employment Department Group and the Office of Population Censuses and Surveys. We identify an occupational sector $O_{i,t}$ for each individual (parents included), for both unemployed and employed individuals. When employed, the occupational sector is defined in a straightforward manner. When unemployed, the occupational sector is interpreted as the sector in which the individual seeks for a job, and is assumed to be the one in which the individual eventually finds a job. For instance, if an individual i starts being unemployed at time t, is unemployed for 10 months and then finds a job in sector 3, we assume that the individual was indeed searching in sector 3 for those 10 months: $O_{i,t} = O_{i,t+1} = ... = O_{i,t+10} = 3$. If we have no information on the occupation of arrival, for instance because the individual exits the sample or goes out of the labor force, we use the occupation prior to the unemployment spell when available. The rationale behind our choices and further details are explained in Appendix A[11]. The distribution of workers across occupational sectors is shown in Table 3.1, along with several labor market statistics of the sectors.

Table 3.1: Distribution of workers across sectors and sectoral labor market statistics. Source: BHPS (1991-2008)

Occupational Sector	**Abs. Freq.**	**Rel. Freq.**	**Unempl. Rate**	**JF Rate**	**JS Rate**
Managers & Administrators	221396	14.88	2.08	10.15	0.25
Professional	147698	9.92	1.82	11.96	0.20
Associate Professional & Technical	174160	11.70	2.59	11.18	0.22
Clerical & Secretarial	235157	15.80	4.35	12.16	0.43
Craft & Related	178628	12.00	6.22	8.10	0.53
Personal & Protective Service	170551	11.46	6.72	8.03	0.52
Sales	108502	7.29	6.39	10.59	0.62
Plant & Machine	131936	8.86	7.82	8.95	0.67
Agriculture & Elementary	120345	8.09	9.66	7.55	0.73
Total	1488373	100.000	4.94	9.38	0.43

We see that labor market outcomes are not independent from sectors. While it is known that high-skilled jobs are better paid, it seems like there are also relatively large differentials in terms of unemployment rate (and search variables). One possible explanation is a relative scarcity of high-skilled workers in the UK in those years, compared to the high profitability of those sectors (managerial and

by other studies, the European economies perform much worse than the US to this extent. For instance, the job finding rate is estimated to be about 30%-40% in the US (Shimer (2012)).

[11] Our results are robust to alternative assumptions: we tried using only the future occupation, the past occupation or a combination of the two.

professional). In the first three sectors the unemployment rate ranges between 1.8% and 2.1%, with a job finding rate of 10-12% and a job separation rate of about 0.2%. On the other hand, we also notice that restricting the sample to the observations for which the occupation is available induces some degree of sample selection. The average job finding rate (unemployment rate) is indeed higher (lower) than in the whole sample. This happens because originally all the individuals assigned to a sector are employed, and our sector imputation only considers the next and the last labor spell. Therefore, due to the large number of missing values for occupation, we lose many observations of unemployed (typically, the long-term unemployed) when imputing the sector. Unfortunately, without making any stronger assumptions than the ones we already make, it is not possible to get rid of this issue. However, notice that the sample selection problem only affects the unemployment rate and the job finding rate, as shown in Table 3.1.

3.3.1 Patterns of Occupational Mobility across Generations

While many studies on occupational mobility across generations rely on single observations for the occupation of parents, the BHPS allows us to follow parents over time. In this way, besides the answer to "What was your parents' occupation when you were 14?", our data provides a better source of information on parents' side.

First of all, we compare the distributions of parents and offspring across sectors. Table 3.2 shows the distribution of sons and daughters, parents when offspring (respondents) were 14 and parents who are followed over time. We immediately notice that a large degree of sex segregation characterizes the distribution across sectors. Managerial and craft occupations are typically covered by men, while secretarial and sales jobs are more intensively taken by women. This phenomenon seems to be persistent over generations, given that no relevant differences can be detected when comparing the distribution of offspring and parents to this extent. Another interesting feature is the large structural change that characterized the UK economy in the last decades. Sectors such as craft or machine occupation shrunk significantly in relative terms, while managerial, professional and especially technical occupations employ nowadays a larger share of the working force than before. For this reason, the distribution of offspring is more directly comparable with the distribution of parents who are followed over time, as in this way we are comparing occupational choices within the same economy.

In what follows, we argue that there exist important differences in occupational mobility computed

Table 3.2: Distribution of sons and parents across sectors, relative frequencies. Source: BHPS (1991-2008)

Occupational Sector	Offspring 1991-2008 Sons	Daughters	Parents when offspring 14 Fathers	Mothers	Parents 1991- Fathers	Mot
Managers & Administrators	18.92	11.96	16.81	7.36	22.48	9.
Professional	10.23	10.60	8.05	6.04	8.69	9.
Associate Professional & Technical	10.83	13.35	4.05	7.61	7.11	9.
Clerical & Secretarial	7.91	24.39	5.00	19.16	6.41	25
Craft & Related	21.12	2.12	27.44	5.40	22.43	2.
Personal & Protective Service	5.90	16.55	5.11	14.90	4.70	18
Sales	4.60	9.81	3.71	11.66	4.36	10
Plant & Machine	13.41	3.29	18.26	7.84	18.35	3.
Agriculture & Elementary	7.07	7.93	11.58	20.03	5.46	11
Total	100.00	100.00	100.00	100.00	100.00	100

using retrospective information and using contemporaneous information, and that these differences are crucial for understanding occupational mobility across generations. In order to investigate the degree of occupational mobility across generations we build Markov matrices, computing the transition probabilities from a sector to another. As parental occupation, we use both the current one and the one as reported when offspring were 14. If individuals rarely switch occupation over the life cycle, the two sources of information on parental occupation will be highly correlated. Consistently with the degree of sex segregation that we found in the data, we report the tables for couples of the same gender: fathers with sons, and mothers with daughters. For males, even though with some heterogeneity, we note that there is a general level of persistence in the same sector as their father's one as reported when respondents were 14. Table 3.3 reveals that the persistence is particularly high at the top (managerial and professional occupations) and at the bottom (plant and machine occupations) of the distribution, with another peak for craft occupations. Instead, when considering the contemporaneous occupation, persistence drops significantly at the top while it strongly increases in the mid-sectors.

When considering women (in Table 3.4), similar considerations can be made: for instance, the persistence with mother's sector as reported when daughters were 14 is strikingly high for managerial and professional occupations. Again, when we look at the contemporaneous occupation, the persistence at the top almost disappears while it becomes more substantial in the middle and at the bottom of the distribution. Overall, women are very attached to clerical and secretarial occupations: the probability of falling into that category is very high regardless of parental background.

Table 3.3: Markov matrix of occupational mobility: fathers-sons, relative frequencies. Source: BHPS (1991-2008)

	Son's sector								
Father's sector when son is 14	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**31.99**	13.76	12.42	6.44	12.56	5.10	4.38	8.30	5.06
2	23.72	**26.11**	21.35	7.55	6.51	4.04	3.38	4.65	2.69
3	25.86	15.75	**17.53**	7.84	12.67	5.16	3.35	8.72	3.13
4	16.98	16.94	13.80	**12.07**	16.71	4.97	3.26	9.69	5.57
5	16.65	9.55	9.32	7.43	**27.86**	5.31	3.76	13.73	6.40
6	18.26	11.61	11.81	10.02	17.87	**8.68**	5.33	11.73	4.68
7	25.09	9.73	13.27	8.88	15.46	2.52	**8.42**	11.55	5.07
8	15.06	7.03	8.38	6.50	24.20	6.50	3.66	**22.13**	6.55
9	14.52	5.47	6.44	7.43	25.18	4.49	4.63	19.37	**12.47**
Father's sector contemporaneous	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**13.46**	7.51	13.11	11.80	22.34	6.94	10.26	5.49	9.09
2	10.77	**13.32**	19.95	16.82	13.02	5.97	10.71	4.46	4.98
3	9.22	7.92	**19.74**	18.96	15.47	5.18	8.36	10.00	5.16
4	13.65	7.85	16.44	**16.03**	17.37	3.82	8.18	10.20	6.46
5	7.99	1.64	8.78	10.95	**40.49**	5.39	7.32	10.28	7.16
6	12.20	1.76	11.69	18.40	17.66	**16.93**	6.87	6.64	7.86
7	7.76	6.47	14.77	13.22	21.95	8.95	**9.95**	11.89	5.03
8	8.95	2.26	6.70	8.50	30.91	7.89	5.51	**20.89**	8.39
9	6.16	4.61	13.16	12.31	23.45	5.35	5.65	8.18	**21.13**

The large differentials in the patterns of persistency obtained by using retrospective instead of current information on parental occupations implies the existence of a substantial degree of occupational mobility of parents over their life cycle. Tables 3.15 and 3.16 (in the Appendix) show that fathers and mothers have a sizeable probability of moving between occupations during their worklife. Studies that focus only on retrospective information on parental occupations cannot account for this important feature of the data.

In order to ease the interpretation of the statistics just shown, we construct likelihood ratios dividing the probabilities shown in the diagonal of the matrix by the unconditional probability that a son (or daughter) belongs to each particular sector. This produces a synthetic measure of the "effect" of parental occupational sector on individual choices. Table 3.5 shows that, when compared to the actual job distributions, the attachment to parental sector is indeed very large. On average, sons (daughters) are 76 % (191 %) more likely to be in the father's (mother's) occupational sector than expected if the assignment of sector was random.

Table 3.4: Markov matrix of occupational mobility: mothers-daughters, relative frequencies. Source: BHPS (1991-2008)

	Daughter's sector								
Mother's sector when daughter is 14	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**18.77**	12.38	15.46	20.77	1.94	15.32	9.03	1.06	5.26
2	13.96	**30.97**	17.75	19.36	0.58	9.26	4.83	0.56	2.73
3	14.72	12.54	**21.03**	21.81	1.28	15.51	6.20	1.95	4.96
4	13.78	16.87	16.28	**29.64**	1.74	11.70	5.51	1.09	3.40
5	12.44	8.30	10.85	22.72	**2.92**	16.39	11.65	4.39	10.33
6	10.86	8.15	13.39	23.85	2.03	**20.11**	9.43	2.57	9.60
7	12.79	8.55	11.93	29.41	1.71	14.61	**11.92**	3.15	5.92
8	10.32	5.02	8.97	26.61	2.76	17.93	10.25	**6.98**	11.15
9	8.89	6.29	13.21	21.58	1.98	20.26	11.50	4.18	**12.12**
Mother's sector contemporaneous	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**16.74**	13.52	9.61	21.04	0.58	19.90	10.83	1.27	6.51
2	13.50	**15.91**	17.27	25.06	1.93	11.80	10.59	0.53	3.41
3	10.01	9.99	**14.89**	25.73	1.30	23.05	9.59	1.35	4.08
4	10.75	7.51	10.45	**37.10**	0.87	18.49	11.01	0.44	3.37
5	6.00	9.25	0.89	23.13	**9.65**	19.29	24.80	3.74	3.25
6	8.07	7.17	10.85	25.41	1.19	**24.59**	16.28	1.15	5.29
7	5.60	3.25	6.02	31.47	2.74	23.57	**18.31**	5.30	3.74
8	6.54	0.94	4.79	23.47	5.87	24.61	13.01	**10.92**	9.84
9	14.12	5.04	9.23	30.27	1.03	19.23	13.98	2.78	**4.31**

Table 3.5: Likelihood ratios: each cell represents the probability that a son (daughter) belongs to a given occupational sector conditional on father (mother) being in the same sector, divided by the unconditional probability of belonging to that sector. Results are shown separately by gender (sons with fathers, and daughters with mothers). Source: BHPS (1991-2008)

Occupational sector contemporaneous	**Likelihood ratio Males**	**Females**
Managers & Administrators	1.29	1.93
Professional	2.59	2.36
Associate Professional & Technical	1.63	1.72
Clerical & Secretarial	1.26	1.55
Craft & Related	1.55	7.26
Personal & Protective Service	1.58	1.44
Sales	1.34	1.73
Plant & Machine	1.94	7.07
Agriculture & Elementary	2.67	1.17

In the next subsection, in order to understand and interpret the patterns of occupational persistence, we study whether parental labor market variables (such as their employment status and their sectoral belonging) affects individuals' labor market outcomes.

3.3.2 Employment Prospects across Generations

Before entering the regression-based analysis, we look at the relationship between labor market performances across generations. In particular, we compute the average unemployment rate and search variables of individuals conditional on the employment status of parents. We also investigate whether these intergenerational correlations vary when individuals are in the same occupational sector as their parents.

Table 3.6: Employment prospects across generations: unemployment rate, JF and JS rate conditional on parental employment status/occupation. Source: BHPS (1991-2008).

		Father			**Mother**	
Variable	**Subsample**	**Unemployed**	**Employed**	**Same Sector**	**Unemployed**	**Employed**
Unemployment Rate	All Sample	20.67	7.84	4.74	17.86	7.86
	Males	24.14	9.11	5.02	20.83	9.18
	Females	14.35	6.23	4.01	15.17	6.28
Job Finding Rate	All Sample	4.85	11.02	15.23	7.45	11.05
	Males	4.88	10.58	14.85	7.65	10.64
	Females	4.86	11.82	16.13	7.20	11.79
Job Separation Rate	All Sample	1.16	0.75	0.58	1.23	0.74
	Males	1.34	0.85	0.58	1.42	0.84
	Females	0.87	0.62	0.57	1.07	0.63

Table 3.6 reveals the existence of strong correlations across generations. Having employed (rather than unemployed) parents is associated with better labor market outcomes. For instance, the average unemployment rate –which is 21% for those whose father is unemployed– drops to 8% for those whose father is employed, decreasing further up to less than 5% when the father is in the same sector as the offspring. Similar percentages characterizes the mother's employment status, with the difference that there does not seem to exist any additional effect linked to sector belonging. The job finding rate more than doubles on average (it increases from 4.9 to 11%) when the father is employed, while the effect of the mother's employment status is less pronounced but still large (from 7.4 to 11%). Again, when the father is employed in the same sector, individuals experience an even higher job finding rate on average (about 15%). Interestingly, the job finding rate of males appears to be affected also by having the

mother in the same sector. Finally, the job separation rate is also correlated with parents' employment status in the same direction. It is roughly 1.1% for those with unemployed father and it drops to 0.75% for those whose father is employed. Mother's employment status has approximately the same effect on this conditional average. An extra reduction in the job separation rate is found when the sector of the offspring coincides with the one of the father, while no relevant differences with respect to the sector of the mother.

Overall, significantly better labor market performances are found to be associated with the employment status of the parents. Such advantages are larger when individuals are in the same occupational sector as their father. The additional premium is about 40-50% the size of the effect of having an employed father[12].

We investigate whether the differences between these groups change over the life cycle. We find that especially for the very young the difference is very large. Figure 3.1 shows that having the father employed is associated with up to 20 to 30 percentage points more in the average employment rate. This difference steadily declines over the life cycle and eventually disappears. The higher employment rate[13] can be generated by higher job finding rates, lower job separation rates or a combination of the two. Figure 3.2 and 3.3 reveal that the job finding rate is driving the bulk of the difference, yielding large and persistent variations across groups. Conversely, the job separation is substantially lower for offspring of employed fathers especially at early ages, whereas the gap greatly reduces later on in the life cycle. Nevertheless, small differences in absolute value are actually large in relative terms and have a strong impact on individual worklife.

The correlations found so far are interesting per se, even though they do not necessarily represent any direct effects of parents on offspring' labor market outcomes. Several other correlations, for instance educational attainment, human capital accumulation or genetical transmission, might well explain these differences in the conditional averages. Moreover, it could also be the case that respondents' outcomes affects parental ones, instead of the other way around. In any case, the differences in the other observables across these groups (Table 3.17 in Appendix 3.A.5) are not large[14]. In Section 3.5 we outline our empirical strategy to address these and other related issues, in order to try to establish a causal relationship and estimate the effect of parental links on offspring' employment prospects.

Before that, we now proceed to postulate a simple of model of intergenerational networks, in order

[12]Remarkably, we do not find that these patterns are substantially different by gender.

[13]It is defined as $1 - unemployment\,rate$.

[14]Not surprisingly, those who belong to the same sector as their father tend to be more often males. Also, they are slighlty older, more educated, more often married and white.

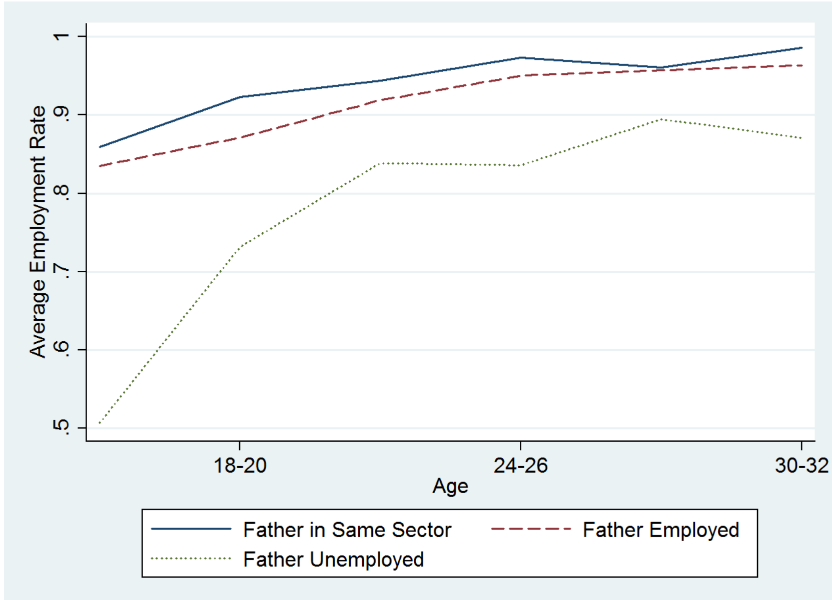

Figure 3.1: Employment Rate (Employed 1, Unemployed 0) as a function of age: cross sectional averages. Source: BHPS.

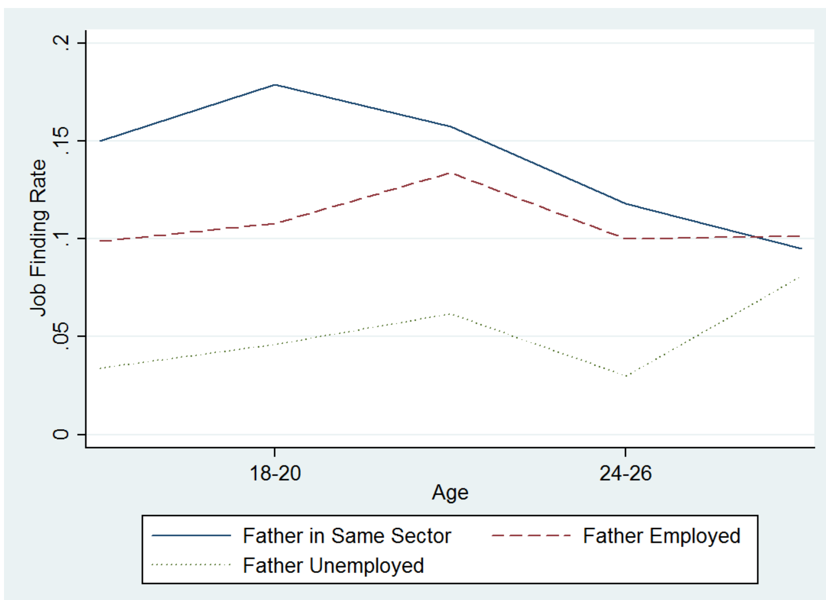

Figure 3.2: Job Finding Rate as a function of age: cross sectional averages. Source: BHPS. Ages 30-32 are cut because of limited availability of observations.

to motivate our empirical strategy and illustrate the source of variation we want to exploit for the identification of a causal relationship.

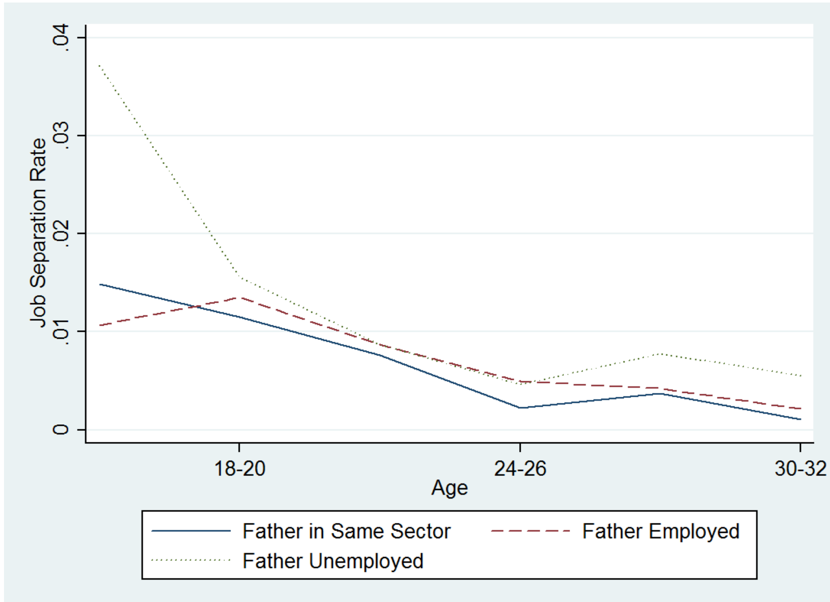

Figure 3.3: Job Separation Rate as a function of age: cross sectional averages. Source: BHPS.

3.4 A Simple Model of Intergenerational Networks

This model is a simple adaptation of Calvó-Armengol & Jackson (2004). Individuals look for jobs when they are unemployed, and in order to do so they exploit their network of social contacts. Employed parents help their unemployed offspring in that they let them use their own social network.

Suppose the economy is populated by identical workers, indexed by worker i, family j and age t. Every period, all workers of age T have an offspring of age 0. We assume that individuals stop being connected with their parents when they have children, so that at each point in time only two generations are connected[15]. In the first period of their lives ($t = 0$), agents draw an initial network size $n^j_{i,0} = \epsilon^j_{i,0}$ from a normal distribution. From that moment onwards, the total network of an offspring ($t < T$) at time t is denoted by $\hat{n}^j_{i,t}$ and it is given by the following expression:

$$\hat{n}^j_{i+1,t} = \beta n^j_{i,t+T} S^j_{i,t+T} + n^j_{i+1,t} \qquad \forall t < T \tag{3.1}$$

where $i+1$ represent the offspring of father i within family j, $n^j_{i,t}$ denotes the natural logarithm of work connections held by worker i in family j at time t and $S^j_{i,t+T}$ denotes the employment status of the father i at age $t+T$.

Fathers' total networks necessarily have to coincide with their personal networks:

[15]This assumption is made for simplicity and does not alter our results.

$$\hat{n}^j_{i,t} = n^j_{i,t} \qquad\qquad \forall t \geq T \tag{3.2}$$

Workers can be in either of two states $S \in \{E, U\}$, employed or unemployed. When employed, they lose their job with constant probability γ. Work connections positively affect the probability of finding a job, as such connections allow workers to reduce informational frictions. Hence we have that, when unemployed, the job finding probability f is:

$$f^j_{i,t} = 1 - e^{-\hat{N}^j_{i,t}} \tag{3.3}$$

where $\hat{N}_{i,t} = e^{\hat{n}_{i,t}}$. We assume that the timing is as follows: first, shocks to the employment status (f and γ) take place; second, personal networks $n^j_{i,t}$ evolve stochastically according to the following law of motion:

$$n^j_{i,t+1} = \begin{cases} \alpha + (1-\delta^E)n^j_{i,t} + \epsilon^j_{i,t} & \text{if } \; S^j_{i,t} = E \\ (1-\delta^U)n^j_{i,t} + \epsilon^j_{i,t} & \text{if } \; S^j_{i,t} = U \end{cases} \tag{3.4}$$

where $\epsilon^j_{i,t} \sim \mathbb{N}(0, \sigma^\epsilon)$. These equations encompass the idea that a worker gains useful connections while working, and may randomly lose/gain more connections every period. While not working, however, such connections depreciate every period because workers progressively lose contact with their former colleagues. In principle the rates of depreciation $\{\delta^E, \delta^U\}$ do not need to be equal, but the difference between them is not important for our results.

It is clear that the correlation between labor market status of fathers and offspring is highest for $t = 0$; at the initial period, connections of offspring are mainly defined by those of their fathers because the former did not have the opportunity yet to form many useful work connections. As time goes by, the careers of fathers and offspring evolve independently and those that were common contacts at the beginning might be still useful contacts for one, but lost touch with the other. As a consequence, the correlation between labor market status fades out along with the correlation between parental and offspring's networks.

Showing formally that the covariance between the employment status of fathers and offspring dies out over time is not straightforward, because the correlation at one point depends on the whole history of employment/unemployment of both the father and the offspring. However, we provide simulation results to show that indeed such correlation fades out as workers get older. These results are shown in Figure

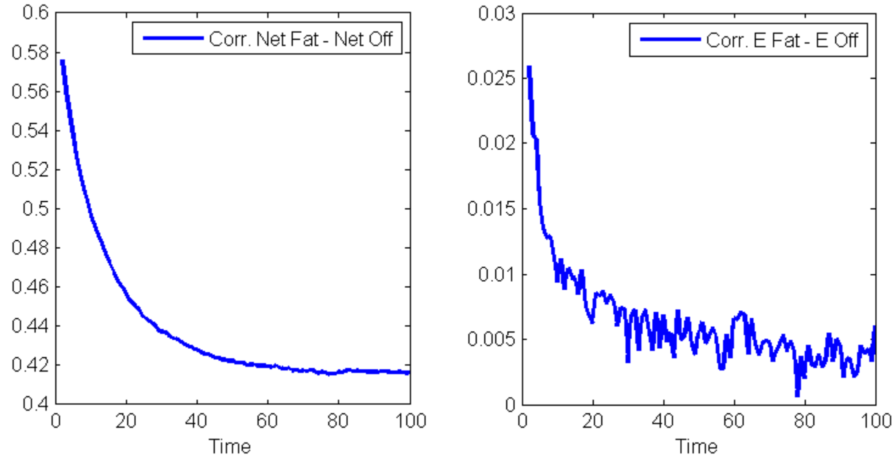

Figure 3.4: Simulated correlation between fathers and offspring **network** (left) and **employment status** (right). $\alpha = 0.05, \beta = 0.5, \delta^E = 0.03, \delta^U = 0.03, \gamma = 0.05$.

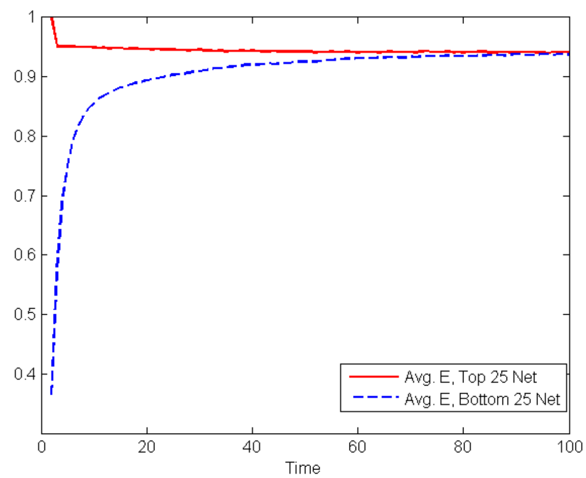

Figure 3.5: Simulated paths of average employment status for individuals with **high** (red line) and **low** (blue line) initial networks. $\alpha = 0.05, \beta = 0.5, \delta^E = 0.03, \delta^U = 0.03, \gamma = 0.05$.

3.4, in which we assign values to the parameters of the model and report the results of our simulations.

Another way to see that differences induced by initial networks vanish over time is to look at the probability of being employed over the life cycle, by different initial conditions. Figure 3.5 shows how individuals with high, rather than low, initial networks have a higher probability of being employed at the beginning of their careers; as time goes by, such difference goes to zero, as we observe in the data.

Shocks to the employment status of the father will have an impact on the employment prospects of offspring mostly at the beginning of their career. This motivates our strategy of looking at the difference in correlation of employment status between ages 20-30 and later ages. As careers evolve independently, the correlation fades out and offspring after age 30 constitute a proper control group for identifying the effect of networks early on.

3.5 Empirical Strategy

We are interested in understanding the partial correlation between individual employment prospects[16] and the employment status of their parents. First of all, we define an employment status variable $E_{i,t}$ using information on job histories. $E_{i,t}$ is equal to 1 if individual i is employed at time t, and 0 in case of unemployment. In all periods of different labor market status (retired, in further education etc.), $E_{i,t}$ is not defined.

Then we define the transition variables $J^f_{i,t}$ and $J^s_{i,t}$, respectively the job finding and job separation events for an individual. $J^f_{i,t}$ is a dummy variable that takes value 1 if individual i moves from unemployment to employment at time t (that is, $E_{i,t-1} = 0, E_{i,t} = 1$), and 0 if the individual remains unemployed ($E_{i,t-1} = E_{i,t} = 0$). In all periods of employment or labor market status different from unemployment, $J^f_{i,t}$ is not defined. Conversely, $J^s_{i,t}$ takes value 1 in case of transitions from employment to unemployment and zero otherwise.

Next, we link individuals and parents using personal identification numbers of relatives provided in the BHPS. For all individuals i for which such information is available, we associate a father, a mother, a spouse and three friends. Call $E^{\textbf{father}}_{i,t}$ the employment status of the father of individual i at time t and similarly for the mother, the spouse and all friends.

In principle we could just use the raw employment status data in our regressions. However, since we have monthly job histories, we are not capable of determining precisely whether jobs ending at time t are covering the full month representing time t or only a small portion of it. The problem is relevant because a correct identification of the timing of spells is crucial to correctly estimate the partial effect of interest: suppose for instance that a father is employed until December 20th when he becomes unemployed, while his offspring obtains a job on December 10th. Since job histories are written in monthly format, it is possible that the father will result unemployed in December, while his offspring will result employed from December onwards. However, it is not clear whether we should have considered the father employed rather than unemployed, since the labor market spell of his offspring began during his employment spell. In order to exclude these controversial cases, we consider only labor market statuses that are unambiguously assigned in a given month, that is we exclude those cases in which the labor market spell

[16] As of now, the focus of our analysis is exclusively on individual employment status and transitions from unemployment to employment (and viceversa). In future work, we are planning to include the wage in our analysis.

changes between two months. Basically, instead of using $E_{i,t}^{\textbf{Father}}$ as defined above, we use

$$E_{i,t}^{\textbf{Father, ongoing}} = \begin{cases} E_{i,t}^{\textbf{Father}} & \text{if } E_{i,t+1}^{\textbf{Father}} = E_{i,t}^{\textbf{Father}} \\ \textit{missing} & \text{if } E_{i,t+1}^{\textbf{Father}} \neq E_{i,t}^{\textbf{Father}} \end{cases}$$

We construct similar variables for mothers and, for comparison purposes, spouses.

3.5.1 Difference-in-Differences Estimation

In order to identify the effect of parental networks on employment prospects, we divide our sample in two groups, one of which is assumed not to be affected by parental networks. Consistently with the stylized model presented in Section 3.4, the control group is made up by all those workers who are not very young anymore. In particular, we employ an age threshold of 27 for discriminating between control and treatment group[17]. The rationale behind this definition of the control group is that individuals accumulate social contacts while working so that their pool of contacts become more and more different from those of their family connections over time. For this reason, an alternative definition of the control group is based on the experience of individuals[18]. For both definitions of control group, we run linear[19] regression models of the form

$$Y_{i,t} = \beta_0 + \beta_1 E_{i,t-1}^{\textbf{Father, ongoing}} + \beta_2 T_{i,t} + \beta_3 T_{i,t} E_{i,t-1}^{\textbf{Father, ongoing}} + \gamma \mathbf{X}_{i,t} + \epsilon_{i,t} \tag{3.5}$$

where $T_{i,t}$ takes the value 1 if the individual belongs to the treatment group (as explained before). The employment status of the father $E_{i,t-1}^{\textbf{Father, ongoing}}$ has a one period lag, in order to avoid problems of double causality (i.e. when the offspring is employed, the father becomes employed thanks to the offspring). $\mathbf{X}_{i,t}$ is a vector of control variables and the dependent variable $Y_{i,t}$ will be, alternatively, the employment status, the job finding rate $J_{i,t}^{f}$ and the job separation rate $J_{i,t}^{s}$. Controls will include a third degree age polynomial, dummies for gender, education, occupational sector (observed for employed, imputed for unemployed), marital status, ethnic group, smoking behaviour, region of residence and

[17]Results are robust to changes in the threshold age. Using any age between 25 and 29 yields a coefficient that yields between 6 and 8 p.p., that is significantly different from zero.

[18]In a robustness check, we define the control group as those workers who have at least a given number of years of potential experience (defined as years elapsed after the completion of education) in the labor market.

[19]In principle, linear models are not ideal for analyses that involve probability because they might predict negative or bigger than one probabilities. We choose linear models over probit/logit formulations because of the easier interpretation of marginal effects. When we run similar logistic regressions, we obtain substantially the same results. Results are now available upon request and will be included in the Appendix of a future version of the paper.

quarterly dummies. We are interested in the estimation of β_3, which will give us the effect of parental networks on labor market outcomes. The identifying assumption is that all other factors affecting the outcome variable other than parental networks affect the offspring of employed and unemployed in the two groups in the same way. That is, we only need that the relative difference in the way these factors affect individuals remains unchanged across the treatment and the control group. Under this assumption, our estimator $\hat{\beta}_3$ will identify the effect we are looking for:

$$\hat{\beta}_3 = (\bar{Y}_{T,E^F=1} - \bar{Y}_{T,E^F=0}) - (\bar{Y}_{C,E^F=1} - \bar{Y}_{C,E^F=0})\,. \tag{3.6}$$

3.5.2 Other Linear Probability Models

In order to check that our results hold when changing the model specification, we also employ three other types of regressions: Pooled Ordinary Least Squares, Random Effects GLS and Fixed Effects. In this case we do not use any control group and our identification strategy with FE estimation estimation crucially depends on the time-invariance of the other factors affecting the outcome variables. The estimating model is a reduced version of the previous one and reads as follows

$$Y_{i,t} = \beta_0 + \beta_1 E_{i,t-1}^{\textbf{Father, ongoing}} + \gamma \mathbf{X}_{i,t} + \epsilon_{i,t} \tag{3.7}$$

We are interested in the estimation of the coefficient β_1. While the POLS is the standard empirical baseline, we are more interested in empirical models that exploit the time structure of the data. In particular, the time-invariant individual heterogeneity captured by the Fixed Effects estimator might affect our results significantly if fixed individual characteristics not captured by controls $\mathbf{X}_{i,t}$ are correlated with labor market outcomes of parents[20]. We run such regressions for both parents and, for comparison purposes, spouses and the three best friends.[21]

For a more in-depth analysis, later on we restrict the sample to those individuals who have employed parents only: that is, an observation is included in the sample if and only if $E_{i,t-1}^{\textbf{Father, ongoing}} = 1$.

Using the occupations $O_{i,t}$, defined for both employed and unemployed individuals as explained in the Data subsection, we compute a new variable $S_{i,t}$, where S stands for "same" sector:

[20]For instance the IQ, motivation, social skills or whatever other factors that are likely to be transmitted across generations and have an impact (directly or indirectly) on the performance in the labor market.

[21]Although we would like to run the same regressions for all these variables at the same time, the low amount of data for which all variables are available does not allow us to do so.

$$S_{i,t} = \begin{cases} 1 & \text{if } O_{i,t} = O_{i,t}^{\textbf{Father}} \\ 0 & \text{if } O_{i,t} \neq O_{i,t}^{\textbf{Father}} \\ \textit{missing} & \text{otherwise} \end{cases}$$

Such variable captures whether an unemployed individual is assumed to be (or not) seeking a job in the same sector of his employed father, or whether an individual is currently working (or not) in the same sector of his employed father.

We run regressions similar to those explained above, where the job finding events $J_{i,t}^{f}$ is regressed on the same sector indicator $S_{i,t}$. Notice that in this case the sample will include only those individual whose parents are employed, meaning that any correlation associated to $S_{i,t}$ will be *additional* to those obtained when looking at the correlation with the employed status of parents.

With the expection of our diff-in-diff specification, in all regressions described above we include individual-level fixed effects, in order to take out fixed individual characteristics that can be correlated with job market outcomes. Moreover, we cluster standard errors at the parents' level, because within-families correlations are likely to be important and to bias standard errors downwards if not properly accounted for.

In the Robustness section we question our empirical strategy, allowing for a more flexible specification; we show that our strategy yields the most "conservative" estimates, and we argue that what seems to be the most "obvious" approach leads to upward biased estimates of the marginal effects. Furthermore, the more flexible specification yields negligible gains in efficiency.

3.6 Results

3.6.1 Difference-in-Differences Estimation

In Table 3.7 we report the results of our diff-in-diff specification. The estimates indicate that having a father employed (rather than unemployed) increases in the individual employment rate of about 8 p.p.. We then decompose this result between an higher job finding rate and a lower job separation, simply by running the same regression changing the dependent variable. Column 2 and 3 of Table 3.7 shows that the bulk of the economic advantage lies in a much higher job finding rate (the effect estimates is 11 p.p.). We want to stress how the effects estimated by our regression are very large and significant. In

	Dependent Variable		
	(1) Emp.Status	(2) Job Finding	(3) Job Separation
Father's emp. status (2m, lagged)	-0.00329	-0.0439	-0.000594
	(0.016)	(0.048)	(0.002)
Younger than 27	-0.0745***	-0.131**	0.00162
	(0.023)	(0.052)	(0.003)
Younger than 27*Father's emp. status (2m, lagged)	0.0822***	0.114**	-0.00253
	(0.023)	(0.049)	(0.003)
N	115823	7912	105727
R^2	0.066	0.040	0.006

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.7: Difference-in-differences regressions of Employment Status, Job Finding and Job Separation. The control group is given by individuals aged at least 27. We report the coefficient of the **employment status of father**, of **belonging to the treatment group** and the **interaction term** (the effect we want to estimate). Standard errors are clustered at the father level. All regressions include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behaviour, marital status, ethnicity, quarter, occupation of search/employment, defined according to the assumptions described in Appendix A.

section 3.6.2 we report the results of other linear probability models in which we control for individual unobserved heterogeneity[22].

Other DD results

The results shown in the last section are perfectly consistent with our stylized model of intergenerational transmission of networks. Importantly, the estimates are robust to alternative definitions of the treatment group. In Table 3.18 (in Appendix 3.A.5), we run again the same regressions using years of potential experience in the labor market (years elapsed from completion of education), instead of age. The control group is defined as all the individuals with more than 10 years of potential experience. The results are virtually unchanged.

We also investigate whether the effect depends on individual characteristics such as age and educa-

[22]If we were to include individual fixed effects in the DD regressions, the identification of the father's effect would rely almost exclusively on cross-sectional variation, in the absence of a large number of individuals who happen to belong to both groups over their life-cycle. In other words, the DD specification would not work properly. On the other hand, the presence of individuals who belong to both the control and the treatment over their life-cycle can also create problems to the identification. Nevertheless, we tried to run the regressions excluding such individuals from the sample and results (available upon request) are substantially unchanged.

tion. In Table 3.19 (in Appendix 3.A.5), we estimate again the baseline regression adding a full set of interaction terms with a gender dummy (in column 1) and a college-education dummy (in column 2). The estimates indicate that the effect of the employments status of the father does not differ by gender, but instead it greatly diminishes when the individual has a college degree. This might reflect the fact that parents are more willing to help their offspring when the latter ones are more disadvantaged (for instance, less educated). Otherwise, this could also mean that the help received by college graduates lies in a better employment, rather than on the margin between employment and unemployment[23].

Finally, we do not find that the same results hold for mothers. Table 3.20 (in Appendix 3.A.5) shows that mothers are not useful work connections for their offspring: in particular, even though the effect on the employment status (column 1) is positive and significant, column 2 reveals that this effect does not arise through higher job finding rates. Therefore, the empirical evidence rejects that having the mother employed help the offspring find a job.

3.6.2 Job Finding Rate - Parental Links

In the remainder of the paper, we focus uniquely on the job finding rate since, as we show in the previous section, the differences in employment prospects of individuals by employment status of the father are mainly driven by differences in finding rates. Table 3.8 shows that having the father employed rather than unemployed has a strong and significant effect on the job finding rate of the offspring, perfectly in line with the results outlined above. The partial correlation observed in POLS models, including all relevant controls, lays in the region of 5-6 p.p. These effects are quite large (to be compared with a 12% in-sample average job finding rate) and robust to several model specifications. Panel regressions with RE yield a similar coefficient (6.4 p.p.). Importantly, the coefficient keeps the same size and it is estimated with a similar precision even in fixed effects models. This suggests that the effects captured by the coefficient do not depend on fixed factors (e.g. genes) that might be transmitted across generations. Notice that the in-sample average job finding rate is higher than the average job finding rate of the overall sample, consistent with the lower average age of the estimating sample. We estimate the baseline POLS regression separately by gender, finding that the father has a large and significant effect both on males and on females.

Conversely, mothers do not appear to have any significant effect on the job finding rate of offspring,

[23] Among college graduates, the unemployment rate is just 3%.

Panel A Dep. Variable: Job Finding

	(1) POLS	(2) POLS (men)	(3) POLS(women)	(4) GLS	(5) FE
Emp. Status (father, lagged)	0.0643***	0.0559***	0.0965***	0.0645***	0.0566***
	(0.014)	(0.018)	(0.028)	(0.016)	(0.021)
Avg. Age (in-sample)	22.1	22.4	21.6	22.1	22.1
Avg. JF rate (in-sample)	0.120	0.114	0.131	0.120	0.120
N	7772	5051	2721	7772	7772
R^2	0.041	0.052	0.061		0.030
N of groups				753	753

Panel B Dep. Variable: Job Finding

	(1) POLS	(2) POLS (men)	(3) POLS (women)	(4) GLS	(5) FE
Emp. Status (mother, lagged)	0.0365**	0.0457*	0.0498	0.0365	0.0266
	(0.018)	(0.025)	(0.035)	(0.022)	(0.026)
Avg. Age (in-sample)	22.6	22.8	22.2	22.6	22.6
Avg. JF rate (in-sample)	0.123	0.118	0.131	0.123	0.123
N	7384	4640	2744	7384	7384
R^2	0.045	0.052	0.078		0.033
N of groups				703	703

Panel C Dep. Variable: Job Finding

	(1) POLS	(2) POLS (men)	(3) POLS (women)	(4) GLS	(5) FE
Emp. Status (Father, lagged)	0.0835***	0.0870***	0.110***	0.0921***	0.0764*
	(0.021)	(0.029)	(0.033)	(0.027)	(0.039)
Emp. Status (Mother, lagged)	-0.000429	-0.0192	0.0195	0.00135	-0.00587
	(0.024)	(0.036)	(0.037)	(0.034)	(0.041)
Avg. Age (in-sample)	22.1	22.3	21.8	22.1	22.1
Avg. JF rate (in-sample)	0.126	0.119	0.137	0.126	0.126
N	5420	3473	1947	5420	5420
R^2	0.047	0.062	0.082		0.037
N of groups				573	573

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.8: Linear regressions of Job Finding Rate (transition from Unemployed to Employed); coefficient for **employment status of father** and **mother** (1 for employed, 0 for unemployed), standard errors (clustered at the father level), average age and average job finding rate in the sample of the regression. Models 1-3 are pooled OLS regressions, model 4 is a random effects GLS regression, model 5 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behaviour, marital status, ethnicity, quarter, occupation of search/employment, defined according to the assumptions described in Appendix A.

neither for males nor for women. The coefficients are ranging between 3 and 4 p.p. but their estimates are less precise, even when we use the data as repeated cross-sections (Table 3.8, panel B).

As the employment status of couples is likely not to be independently distributed, our models might be suffering from omitted variable bias. In order to control for correlations between the employment status of the father and of the mother, we estimate the model including both regressors. The results shown in Table 3.8 (panel C) confirm the patterns shown in the separate regressions, yielding the father's employment status as the only important predictor of offspring's transitions. This is consistent with other studies as for instance Magruder (2010). The effect of having the father employed ranges between 7.6 and 11 p.p., while the effect of the mother is not stable across specifications and never significantly different from zero. This suggests that the positive effects of mother's employment status –shown in panel B of Table 3.8– were almost entirely driven by within-couple correlation in employment status. Notice that, even though standard errors rise in fixed effects estimation, the father's coefficient keeps having the same size (or even higher). This indicates that such effects do not depend on within-household correlation in employment status.

In order to get further insights on the father effects found so far, we test whether these are magnified when the occupational sectors of the offspring and of the father coincide. That is, we investigate whether individuals who search for a job in the same sector where their father is employed have additional advantages. As shown in Table 3.9, such additional advantages are estimated to be in the region of 4 p.p.[24]. The size of the coefficient is again robust to the inclusion of individual fixed effects. This is a substantial difference and it might be one of the main factors driving the occupational persistence across generations that we find in the data.

We do not find any effects for mothers, consistently with our previous findings. Mothers' employment status does not appear to provide any advantages to offspring, not even when their job is similar to the one their offspring are looking for (the associated regression table is available upon request).

[24] This implies that having the father employed in the same sector where individuals are looking for jobs generates an effect that is at least 1.6 larger than having the father employed in some other sector. The effect found in Table 3.8 is a composition of the effect of fathers in the same sector and fathers in other sectors. For this reason, computing the additional effect as the ratio between the two estimates (0.4/0.6=0.66..) simply provides a lower bound.

	Dep. Variable: Job Finding		
	(1) POLS	(2) GLS	(3) FE
Emp. in Same Sector (father)	0.0422**	0.0447**	0.0407
	(0.018)	(0.021)	(0.026)
Avg. Age (in-sample)	22.0	22.0	22.0
Avg. JF rate (in-sample)	0.130	0.130	0.130
N	6257	6257	6257
R^2	0.045		0.031
N of groups		666	666

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.9: Regressions of Job Finding; coefficient for **father in same sector** (0 employed in other sector, 1 employed in same sector), standard errors (clustered at the father level), average age and average job finding rate in the sample of the regression. Model 1 is a pooled OLS regression, model 2 is a random effects GLS regression, model 3 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behaviour, marital status, ethnicity, quarter, occupation of search/employment, defined according to the assumptions described in Appendix A.

3.6.3 Job Finding Rate - Comparison with other strong Links

In this section we consider the employment status of the three closest friends and of the spouse, investigating whether these partial correlations are similar in magnitude to the parental ones we documented in the previous section. To ease the comparison we employ the same empirical strategy and model specifications. The only difference is that we do not distinguish between males and females in the regressions. In the first model, we consider the number of employed friends[25], among the three closest as reported by individuals. Table 3.10 reveals that friends' employment status has a significant impact on the probability of transition from unemployment to employment. Having an additional employed (rather than unemployed) friend raises on average the individual job finding rate by 3 p.p. Notice that this coefficient is significantly smaller than the father's coefficient (about half in magnitude). Moreover, the friends' coefficient drops with the inclusion of fixed effects in the model, suggesting that individual characteristics are producing a bias in the baseline regressions. Some fixed factors are positively correlated with both the ability of finding a job and having good (employed) friends. Our estimates are in line with those of Cappellari et al. (2010), who find an effect of about 7.4 p.p. on yearly transitions.

[25]We follow the same strategy as Cappellari & Tatsiramos (2011).

Panel A	Dep. Variable: Job Finding		
	(1)	(2)	(3)
	POLS	GLS	FE
Num. Employed Friends (lagged)	0.0287***	0.0275***	0.0110
	(0.003)	(0.005)	(0.008)
Avg. Age (in-sample)	33.5	33.5	33.5
Avg. JF rate (in-sample)	0.101	0.101	0.101
N	14127	14127	14127
R^2	0.028		0.031
N of groups		1919	1919

Panel B	Dep. Variable: Job Finding		
	(1)	(2)	(3)
	POLS	GLS	FE
Emp. Status (spouse, lagged)	0.0527***	0.0690***	0.0614***
	(0.010)	(0.015)	(0.022)
Avg. Age (in-sample)	43.3	43.3	43.3
Avg. JF rate (in-sample)	0.100	0.100	0.100
N	10580	10580	10580
R^2	0.027		0.021
N of groups		1075	1075

Standard errors in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 3.10: Regressions of Job Finding Rate (transition from Unemployed to Employed); coefficient for **number of employed friends** (from 0 to 3) **employment status of spouse** (0 or 1), standard errors (clustered at the individual level), average age and average job finding rate in the sample of the regression. Model 1 is a pooled OLS regression, model 2 is a random effects GLS regression, model 3 is a fixed effects regression. All models include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behaviour, marital status, ethnicity, quarter, occupation of search/employment, defined according to the assumptions described in Appendix A.

As Table 3.10 shows, spouse links seem to be stronger than those of friends. According to regression results, individuals whose spouse is employed experience a job finding rate that is 5-6 p.p. higher than that of individuals who are married to an unemployed spouse. One possible concern is assortative mating, i.e. the fact that people who are more likely to be employed tend to marry among them. However, the fact that the size of the estimated coefficient is robust to fixed effects estimation strategies suggests that this mechanism is not driving the results. Summing up, spouse effects are comparable in size to father's ones, while friendship ties seem to be a less important factor in the determination of the job finding rate.

3.7 Discussion

In this section we discuss our results and provide some possible alternative mechanisms that can explain the partial correlations observed in the data. The focus of our discussion is exclusively on the effects of fathers' variables on offspring' job finding rates, which we consider as the most important of our results. At a first glance, these positive effects are consistent with standard models of networks in the labor market (Calvó-Armengol & Jackson (2004)). Information flow on vacancies and job opportunities probably represents one of the main channels through which individuals belonging to the same social network help each other. Of course, the partial correlations uncovered by our regressions possibly include several other mechanisms.

Genetical and Human Capital Transmission

For instance, genetical and human capital transmission across generations might be driving the results. To this respect, we have to consider that for each of the models we estimate, we always include fixed effects as the last specification. In this way we capture fixed individual characteristics that have an effect on the dependent variable and are possibly correlated with the explanatory variables of interest [26]. Genetical endowments are an example of such individual characteristics that are properly controlled for in fixed effects models, assuming that their effect is linear and non time-varying. With respect to human capital, even though it could –at least in part– be assimilated to fixed factors in adult individuals, this is certainly not true for young individuals. Human capital is a time-varying factor that can be potentially relevant in our estimates. The presence of educational group dummies in our regression attenuates this problem, as education is a good proxy for human capital. However, if the effects were due to the transmission of human capital or work ethics, then we should find that the exact timing of the employment status (or the sectoral belonging) of the father does not matter much. Indeed, such transmission mechanisms are supposed to be long-lasting, and it is also reasonable to think that they take some time in order to produce their effects. Hence, as a further robustness check we include in our

[26] To some extent, father's fixed effect might be better at capturing fixed characteristics that are transmitted across generations. Including such fixed effects in our regression -rather than individual ones- leaves the results unchanged.

regression several lags of the employment status of the father. Interestingly, columns 1-4 of Table 3.11 reveals that only the contemporaneous employment status and sector of the father have an effect. The coefficient of the lags considered (3, 6 and 12 months) are actually negative or not significant, indicating that human capital transmission is not a relevant factor in our estimates. Since strong collinearity might be causing a bias in our coefficients, we also estimate a regression including only the 12-months lag of our variables of interest (columns 2 and 4 of Table 3.11). We find this to have, if anything, a slightly negative effect on the job finding rate. Repeating the same test for both the employment status and the occupational sector provides a test for, respectively, a general and an occupation-specific human capital interpretation. As we can see, the empirical evidence is strongly at odds with this interpretation. The fact that the coefficients of the current variables are even higher now reveals that in the baseline models these coefficients were picking up the negative correlations of the lagged variables, which are serially correlated.

Direct Hiring

Second, another possible channel is direct hiring of individuals by their father. Even though it is unclear whether this should be considered as an informational advantage or another kind of mechanism, we investigate whether a major part of the effects we find can be attributed to this channel. We study whether having a father who hires employees (rather than employee or self-employed without employees) boosts the advantages in terms of job finding rate. Column 5 of Table 3.11 shows that, if anything, having a father who is an employer has a negative effect on the individual job finding rate. This is strongly inconsistent with an interpretation of our results as direct hiring.

Local Labor Market Conditions

Another possibility is the existence of common shocks affecting both parental employment status and offspring' performances. For instance, if an individual and his father both live in a region that has experienced a positive shock, their employment statuses will be correlated as they will be caused by the same fundamental shock. Similar considerations can be made with respect to the occupational sector. If the partial correlations we find are due to local labor market conditions, then we should expect these correlations to be stronger when the offspring lives together with his father. To this purpose, we use

	Dep. Variable: Job Finding				
	(1)	(2)	(3)	(4)	(5)
	Lags	Lag 12 only	Sector lags	Sector lag 12 only	Fat. Hires Employees
Emp. Status (father, lagged)	0.0759*				0.0577***
	(0.041)				(0.019)
Emp. Status (father, 3 months lag)	-0.0275				
	(0.036)				
Emp. Status (father, 6 months lag)	-0.0238				
	(0.030)				
Emp. Status (father, 12 months lag)	0.00726	-0.0257			
	(0.025)	(0.024)			
Father in Same Sec. (lagged)			0.0790**		
			(0.040)		
Father in Same Sec. (3 mths lag)			0.00338		
			(0.051)		
Father in Same Sec. (6 mths lag)			-0.0197		
			(0.040)		
Father in Same Sec. (12 mths lag)			0.00691	-0.0117	
			(0.028)	(0.021)	
Father Hires Employees					-0.106
					(0.082)
N	6621	7648	4841	5855	8563
R^2	0.026	0.024	0.035	0.029	0.027
N of Groups	654	719	554	624	791

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.11: Discussion: Human Capital and Direct Hiring. All regressions are fixed effects estimates. All regressions include all controls discussed in previous sections. Standard errors are clustered at the father level.

the region of residence, generating a dummy that takes the value 1 when the region of residence of the offspring does not coincide with the one of the father. Column 1 of Table 3.12 shows that the partial correlation of father's employment with offspring' job finding rate is instead magnified when the offspring lives in a different region from his father, even though the estimate of the difference is not very precise. In our regression we control for regional changes of offspring, to account for the possibility that individuals migrate in order to find a job, which would bias the estimate of the father's employment coefficients. Individuals who belong to different regions definitely belong to different local labor markets, and therefore we have to conclude that local labor market conditions are not an important driver of the correlations we find. In order to control further for local labor market conditions, we also compute the average unemployment rate by sector and by region. We then add these new variables to our regressions

Dep. Variable: Job Finding

	(1) Local Conditions	(2) Sector Unemp.	(3) Region Unemp.	(4) Sector*Year	(5) Region*Year
Emp. Status (father, lagged)	0.0558***	0.0553***	0.0553***	0.0564***	0.0542**
	(0.021)	(0.019)	(0.020)	(0.021)	(0.021)
Father emp. (in other region)	0.0619				
	(0.049)				
Father unemp. (in other region)	-0.121				
	(0.150)				
Has Changed Region from last year	X				
Unemployment of Sector		X			
Unemployment in Metropolitan Area of residence			X		
Interactions Sector × Year				X	
Interactions Region × Year					X
N	7816	8563	8563	8563	9246
R^2	0.029	0.027	0.029	0.047	0.062
N of Groups	754	791	791	791	828

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.12: Discussion: Local Labor Market Effects. All regressions are fixed effects estimates. All regressions include all controls discussed in previous sections. Standard errors are clustered at the father level.

as additional controls. As shown in columns 2-5 of Table 3.12, the partial correlations are unchanged by the inclusion of all these possible controls. In particular, we are including dummies for the sector interacted with the year in column 4 and for the region interacted with the year in column 5, controlling for possible booms or busts of given segments of the labor market. Nonetheless, this does not appear to capture at all the effects outlined so far.

3.8 Robustness Checks

In this section we explore whether our results are robust to different choices of the sample and to different empirical strategies. First, we want to understand whether the composition of our sample might be driving our estimates. The fact that our estimating sample includes many individuals who are still at school or at the university might be creating problems of sample selection. To control for this possibility, we try to exclude individuals with a college degree from our sample. Column 1 of table

3.13 presents results of this estimation: although the size of the estimate is somewhat lowered, it still is statistically and economically significant, showing that college-educated individuals are not driving the bulk of the correlations we find.

Then, we consider the possibility that only very young workers (aged 16-20) are affected by the employment status of their father. However, when we include only individuals aged more than 20 years of age in the estimation (column 4), we maintain the size of the coefficient, despite losing more than one-third of our original sample.

Also, we consider the possibility that our assumptions on sectors of search might be important for our results: by using future and past occupations as proxies of current sectors of search, we are de facto excluding those individuals who are always unemployed in the BHPS, or who never report their occupation. To account for this possibility, we exclude controls for occupation from our estimations. Results are reported in column 2: the coefficient is lowered by about 1 percentage point, maintaining statistical and economical significance.

Finally, we question our empirical strategy and consider the possibility that a more flexible model may allow us to better capture the nature of the correlations we find. That is, we do not keep only fathers who are on an ongoing spell but rather all the observations which are not missing. Specifically, we construct four indicators based on the two months of job history of the father during the offspring's transitions: hence we have one dummy for "father unemployed past month and current", one for "father unemployed past month but employed on current" and so on. Column 3 shows the results of such experiment: while the coefficient roughly corresponding to our empirical strategy (employed past month and current) maintains substantially the same magnitude and standard error, the coefficient corresponding to "father unemployed last month, employed today" is strikingly high. Such a coefficient is due to a relatively large number of transitions taking place at the same time (for both fathers and offspring) and does not correctly capture any direct effect of fathers on offspring. There are at least two main issues: first, common high-frequency shocks that we are not able to properly control for might be a common cause for these contemporaneous events. Second, there is the possibility that in fact offspring are affecting fathers (instead of the other way around), producing a large upward bias due to reverse causality.

	(1) No College	(2) No Sectors	(3) Different Model	(4) Age > 20
Emp. Status (father, lagged)	0.0461***	0.0490***		0.0616**
	(0.017)	(0.017)		(0.025)
Father U-E			0.172**	
			(0.072)	
Father E-U			0.0526	
			(0.057)	
Father E-E			0.0632***	
			(0.019)	
N	7826	9246	8644	5889
R^2	0.026	0.024	0.026	0.031
N of groups	671	828	792	576

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.13: Robustness Checks: regression without college graduates (column 1), no sectoral dummies (column 2), more flexible specification (column 3), Only individuals aged > 20 (column 4). Omitted category: Father U-U. All regressions are fixed effects estimates. All regressions include all controls discussed in previous sections. Standard errors are clustered at the father level.

3.9 Conclusion

We tested whether parental links affect labor market outcomes of individuals using rich panel data from the British Household Panel Survey. Our results indicate that, on average, those whose father is employed rather than unemployed experience an employment rate that is about 8 percentage points higher, with job finding rates which are higher by 5 percentage points and job separation rates which are lower by 0.3 p.p.. We also show that such difference is larger when individuals work in occupations similar to those of their father. We do not find similar correlations for mothers, and we show that father's effects are similar in magnitude, or larger, to those of other supposedly relevant links. We also document that the job separation rate is on average lower for individuals whose father is employed in similar occupations to theirs.

By means of a number of robustness checks, we show that our results are unlikely to be attributable to human capital transmission, to common shocks driving both outcomes at the same time or to the fact that fathers directly hire their offspring. Our conclusion is that parental networks are likely to play an important role in determining labor market outcomes.

Appendix

3.A The Data

3.A.1 Construction of Job Histories

The BHPS is a yearly survey, and therefore its basic structure contains yearly observations for each individual. Among the available variables, individuals report what their employment status and occupation is at the moment of the interview and when the current spell began. In addition to the main dataset, there is a separate annex in which individuals list their detailed job history in the last 12 months. Each single spell is identified with a start date and an end date. When the month of the start date is missing, we replace it with the month of the interview (if the spell began in the same year) or with December (if the spell began in some previous year). In this way, we partly exploit those spells, that otherwise would be completely missing. For each spell we are provided with the employment status, occupation and other information.

We replace the yearly observations by 12 monthly observations for each individual. Then, we fill in the employment status exploiting the information provided. Constructing correctly the job histories is not a straightforward exercise, as the spells reported by individuals sometimes overlap or conflict with each other. In order to solve this issue, we need to set a hierarchical order of the available information. Importantly, we never replace the variables we copy over time once they are assigned, even if they get into conflict with some future source of information. We give priority to the current spell report, as the amount of recall needed to report it correctly is smaller than for past spells. Therefore, first of all we copy the current employment status over time, from the start date of the current spell to the date of the interview. Second, we use past spells to fill in the remaining missing values. Again, we assume that recalls closer in time are more reliable and therefore we first consider the very last spell, then the second last and so on.

We fix 12 as the maximum number of difference in months between the interview (moment of the recall) and the variable to assign (object of the recall). For individuals who are interviewed every year this choice has virtually no effect, as their employment sequences are constructed simply using for each given year the information provided in the interview of the same year. For the others, this choice is meant to limit the amount of measurement error generated by imperfect recall. We noticed that individuals often change their answer to the lenght of the current spell or modify the order or the nature of a job spell, even after years. This implies that without fixing a maximum time difference for assigning the variables, we would end up with a dataset that included pieces of different spell, oftern misreported, one after the other.

3.A.2 Employment Status Imputation of Friends

Individuals are asked about the employment status of their friends once per two years. What is available in the basic structure of the BHPS is therefore a unique observation. Unfortunately we cannot construct the job histories of friends, as no identification number is reported. In order to keep the monthly frequency, we replicate the information on friends over the following 12 months. This is done also to keep relatively large the sample size. Our imputation procedure is based on the assumption that the employment status features a relatively large degree of persistence over time. This is certainly true for employment spell, as the job separation rate in the sample is small and implies long average job duration. It is also true for unemployment spells, as the average unemployment spell duration is above one year. By replicating the employment status in the following 12 months we are simply assuming that those spells of friends are average ones. The only risk we bear is to misplace them in time.

3.A.3 Sector Imputation of Unemployed

The unemployed, by definition, do not belong to any occupational sector. One might even argue that unemployed are simply looking for some job, regardless of any occupational classification. Instead, we believe that we gain useful insights by imputing sector of search to the unemployed. From the data we see that individuals do not change occupational sector often and, even when they do so, the change is usually not dramatic (e.g. movements from sector 2 to sector 3). Moreover, it seems reasonable to think that individuals target their job search to some particular sector of the economy, consistently with their educational level, qualifications and past occupations. Therefore we treat unemployed workers -for which the sector is in principle missing- as if they were still belonging to some occupational sector.

Furthermore, for the purpose of our analysis we need to assign them to some sector.

The problem is that we do not really know in which sector they are seeking jobs. The idea behind our imputation is very simple: by logic, the sector where an unemployed worker finds a job is just the sector where he was seeking jobs. The only limitation is that we assign the whole unemployment spell to that particular sector, without allowing for movements across sectors within the spell. When the sector after the unemployment spell is not reported, then we use the previous sector. In any case, to limit the amount of measurement error generated by our imputation, we only consider spells that immediately follow (or precede) the unemployment spell of interest.

3.A.4 Educational and Occupational Classification

For constructing educational groups, we consider the highest educational qualification achieved by individuals. The original variable contains more than ten possible values, with an elevated degree of details. We collpase those ten groups into four. The first group corresponds to those who hold a Bachelor's degree or some higher degree. The second group includes the individuals with a high school diploma or qualifications for teaching or nursing. Individuals with an A level or O level fall into the third group. Finally, the fourth group is for those who hold no qualification whatsoever.

With respect to the occupational classification, we follow the aggregation in major group of the SOC as proposed by the Employment Department Group and the Office of Population Censuses and Surveys. The BHPS uses the SOC90 (Standard Occupational Classification), a three-digit code, for describing occupations. At the most disaggregated level we have 347 categories, and in order to analyze the persistence across sectors we need to aggregate them. We choose the aggregation in major groups (9 categories) as the one able to preserve some substantial degree of persistence while keeping a satisfactory level of details.

For further details, refer to "Standard Occupational Classification - Structure and Definition of Major, Minor and Unit Groups, Volume 1"

3.A.5 Representativeness of the BHPS sample

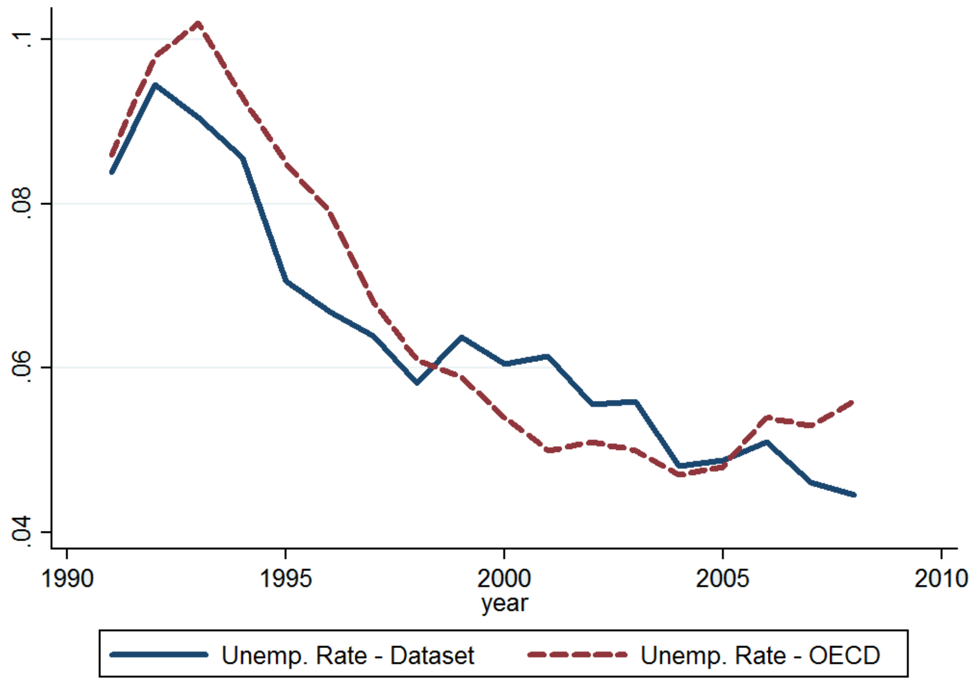

Figure 3.6: In-sample unemployment rate compared to the Harmonized Unemployment Rate in UK, 1991-2008 (Source: OECD).

Appendix B - Additional Tables

Summary Statistics

Table 3.14: Summary statistics of labor market outcomes. Source: BHPS (1991-2008).

Variable	Subsample	Mean	N
Unemployment Rate	All Sample	6.15	1685930
	Males	7.24	865469
	Females	5.00	820051
Job Finding Rate	All Sample	7.21	99532
	Males	7.15	60356
	Females	7.32	39121
Job Separation Rate	All Sample	0.42	1549143
	Males	0.49	786531
	Females	0.34	762282

3.A.6 Markov Matrices

Table 3.15: Markov matrix of occupational mobility: fathers when sons are 14-fathers over their life-cycle, relative frequencies. Source: BHPS (1991-2008)

Father's sector when son is 14	1	2	3	4	5	6	7	8	9
1	**59.32**	4.41	6.16	7.32	6.31	1.63	6.33	7.10	1.42
2	16.03	**68.07**	9.13	3.07	0.72	0.33	1.12	0.66	0.86
3	11.72	11.11	**54.46**	10.82	4.28	0.18	4.85	1.01	1.58
4	14.72	2.10	2.60	**49.86**	3.38	8.08	15.22	4.04	0.00
5	5.16	2.54	3.89	1.85	**65.88**	3.19	0.37	12.75	4.37
6	9.54	0.19	18.41	21.96	3.61	**38.66**	1.94	3.51	2.18
7	27.79	0.00	1.59	9.95	4.58	0.00	**33.60**	18.49	4.01
8	8.24	3.47	1.08	1.11	8.91	1.86	1.15	**70.11**	4.07
9	14.81	15.68	0.50	4.16	12.28	5.45	4.18	22.36	**20.59**

Table 3.16: Markov matrix of occupational mobility: mothers when daughters are 14-mothers over their life-cycle, relative fequencies. Source: BHPS (1991-2008)

Mother's sector when daughter is 14	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**
1	**46.01**	0.00	6.78	21.68	0.00	16.29	1.54	6.36	1.33
2	0.98	**62.87**	19.81	7.57	0.00	0.94	0.00	2.29	5.53
3	1.61	7.20	**69.76**	4.66	0.72	10.90	0.76	0.68	3.70
4	17.05	3.21	5.88	**62.81**	0.88	2.11	5.04	0.27	2.75
5	2.25	0.00	3.56	43.34	**12.66**	9.10	16.60	6.94	5.53
6	13.49	2.59	6.14	12.88	0.23	**46.58**	6.44	1.68	9.97
7	7.61	5.80	1.91	12.83	4.92	13.02	**40.40**	0.81	12.70
8	3.02	0.98	13.06	28.24	9.71	3.59	8.00	**28.41**	4.98
9	7.08	4.60	5.21	6.17	0.28	12.29	12.83	3.19	**48.34**

Descriptive Statistics of different groups

Table 3.17: Descriptive statistics (averages) of different groups: offspring of unemployed fathers, employed fathers and employed fathers in their same occupational group. Source: BHPS (1991-2008).

Variable	**Father unemployed**	**Father Employed**	**Father in Same Sector**
% Female	45	48	32
% Smoker	27	25	30
Age	21.44	22.22	23.49
% College-educated	6	12	13
% Married	6	10	13
% Non-White	7	3	1
Modal Sector	6	4	5

Diff-in-diff: Alternative Definition of Control Group

	Dependent Variable		
	(1)	(2)	(3)
	Emp.Status	Job Finding	Job Separation
Father's emp. status (2m, lagged)	0.00415	-0.0254	-0.000894
	(0.025)	(0.037)	(0.003)
Less than 11 yrs. of pot. experience	-0.0563	-0.0876**	0.000791
	(0.038)	(0.043)	(0.003)
Less than 11 yrs. of pot. experience* Father's emp. status (2m, lagged)	0.0770**	0.103***	-0.00102
	(0.036)	(0.038)	(0.003)
N	117110	8160	106833
R^2	0.069	0.041	0.007

Standard errors in parentheses
* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.18: Difference-in-differences regressions of Employment Status. The control group is given by individuals with more than 10 years of potential experience in the labor market. We report the coefficient of the **employment status of father**, of **belonging to the treatment group**, the **interaction term** (the effect we want to estimate). Standard errors are clustered at the father level. All regressions include a third-degree polynomial in age and dummies for education, gender, region of residence, smoking behaviour, marital status, ethnicity, quarter, occupation of search/employment, defined according to the assumptions described in Appendix A.

Diff-in-diff: Heterogeneity Analysis

	Dependent Variable	
	(1)	(2)
	Emp.Status	Emp.Status
Father's emp. status (2m, lagged)	0.0120	0.0155
	(0.032)	(0.026)
Younger than 27	-0.0834	-0.0613
	(0.051)	(0.040)
Younger than 27*Father's emp. status (2m, lagged)	0.0856*	0.0802**
	(0.051)	(0.039)
Younger than 27*Father's emp. status (2m, lagged)*Female	-0.0307	
	(0.066)	
Younger than 27*Father's emp. status (2m, lagged)*College		-0.105**
		(0.052)
N	120460	120460
R^2	0.068	0.064

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.19: Difference-in-differences regressions of Employment Status. The control group is given by individuals aged at least 27. We report the coefficient of the **employment status of father**, of **belonging to the treatment group**, the **treatment effect** and the **treatment effect interacted** with Female in Column 1 and with College in Column 2 (the differential effect we want to estimate). Interactions between Female (or College, respectively) and the Employment Status of Father and the treatment group indicator are also included in the model, altough the coefficients are not reported. Standard errors are clustered at the father level. All regressions include all controls previously discussed.

Diff-in-diff: Regressions for Mother

	Dependent Variable		
	(1)	(2)	(3)
	Emp.Status	Job Finding	Job Separation
Mother's emp. status (2m, lagged)	-0.00894	0.0664***	0.00394*
	(0.027)	(0.024)	(0.002)
Younger than 27	-0.0779**	0.0283	0.00624
	(0.040)	(0.038)	(0.005)
Younger than 27*Mother's emp. status (2m, lagged)	0.0835**	-0.0376	-0.00753
	(0.039)	(0.032)	(0.005)
N	110711	7728	100897
R^2	0.060	0.044	0.006

Standard errors in parentheses

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 3.20: Difference-in-differences regressions of Employment Status, Job Finding and Job Separation. The control group is given by individuals aged at least 27. We report the coefficient of the **employment status of mother**, of **belonging to the treatment group**, the **interaction term** (the effect we want to estimate). Standard errors are clustered at the mother level. All regressions include all controls previously discussed.

Bibliography

ADDISON, JOHN T., & PORTUGAL, PEDRO. 1989. Job Displacement, Relative Wage Changes and the Duration of Unemployment. *Journal of Labor Economics*, **7**(3).

AGUIAR, MARK, & HURST, ERIK. 2007. Measuring Trends in Leisure: The Allocation of Time over Five Decades. *Quarterly Journal of Economics*, **122**(3).

ALLARD, MARY DORINDA, BIANCHI, SUZANNE M., STEWART, JAY, & WIGHT, VANESSA R. 2004. Inequality in Parental Investment in Child-Rearing: Expenditures, Time, and Health. *In:* NECKERMAN, KATHRYN M. (ed), *Social Inequality*. New York: Russel Sage Foundation.

ALLARD, MARY DORINDA, BIANCHI, SUZANNE M., STEWART, JAY, & WIGHT, VANESSA R. 2007. Comparing Childcare measures in the ATUS and in Earlier Time-Diary Studies. *Monthly Labor Review*, **130**(5).

ANGER, SILKE, & HEINECK, GUIDO. 2009. Do Smart Parents Raise Smart Children? The Intergenerational Transmission of Cognitive Abilities. *Journal of Population Economics*, **23**(3).

ATKINSON, ANTHONY B. 1970. On the Measurement of Inequality. *Journal of Economic Theory*, **2**.

BARNETT, JUSTIN, & MICHAUD, AMANDA. 2012. Wage Scars from Job Loss. *Working Paper*.

BAYER, PATRICK, ROSS, STEPHEN, & TOPA, GIORGIO. 2008. Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes. *Journal of Political Economy*, **116**(6).

BEAMAN, LORI A. 2012. Social Networks and the Dynamics of Labour Market Outcomes: Evidence from Refugees Resettled in the U.S. *Review of Economic Studies*, **79**(1).

BECKER, GARY, & TOMES, NIGEL. 1986. Human Capital and the Rise and Fall of Families. *Journal of Labor Economics*, **4**(3).

Becker, Gary S., & Tomes, Nigel. 1976. Child Endowments and the Quantity and Quality of Children. *Journal of Political Economy*, **84**(4).

Becker, Gary S., & Tomes, Nigel. 1979. An Equilibrium Theory of the Distribution of Income and Intergenerational Mobility. *Journal of Political Economy*, **87**(6).

Bentolila, Samuel, Michelacci, Claudio, , & Suarez, Javier. 2010. Social Contacts and Occupational Choice. *Economica*, **77**(305).

Bernal, Raquel, & Keane, Michael P. 2010. Quasi-Structural Estimation of a Model of Child Care Choices and Child Cognitive Ability Production. *Journal of Econometrics*, **156**.

Bernal, Raquel, & Keane, Michael P. 2011. Child Care Choices and Children's Cognitive Achievement: The Case of Single Mothers. *Journal of Labor Economics*, **29**(3).

Bianchi, Suzanne M. 2000. Maternal Employment and Time with Children: Dramatic Change or Surprising Continuity? *Demography*, **37**(4).

Björklund, Anders, Roine, Jesper, & Waldenström, Daniel. 2012. Intergenerational top income mobility in Sweden: Capitalist dynasties in the land of equal opportunity? *Journal of Public Economics*, **96**(5-6).

Black, Sandra E., Devereux, Paul J., & Salvanes, Kjell G. 2008. Like father, like son? A note on the intergenerational transmission of IQ. *IZA Discussion Papers*.

Boca, Daniela Del, Flinn, Christopher, & Wiswall, Matthew. 2012. Household Choices and Child Development. *Working Paper*.

Brilli, Ylenia. 2012. Mother or Market Care? A Structural Estimation of Child Care Impacts on Child Development. *Working Paper*.

Brooks, Bradley. 2005. Chronic Unemployment: a Statistical Profile. *Statistics Canade, Analysis in Brief*, **31**.

Calvó-Armengol, Antoni, & Jackson, Matthew O. 2004. The Effects of Social Networks on Employment and Inequality. *The American Economic Review*, **94**(3).

Calvó-Armengol, Antoni, & Zenou, Yves. 2005. Job Matching, Social Network and Word-of-Mouth Communication. *Journal of Urban Economics*, **57**(3).

Cappellari, Lorenzo, & Tatsiramos, Konstantinos. 2011. Friends' Networks and Job Finding Rates. *Working Paper*.

Carlson, Marcia J., & Corcoran, Mary E. 2001. Family Structure and Children's Behavioral and Cognitive Outcomes. *Journal of Marriage and Family*.

Chèron, Arnaud, Hairault, Jean-Olivier, & Langot, FranÃ§oise. 2013. Life-Cycle Equilibrium Unemployment. *Journal of Labor Economics*, **31**(4).

Clark, Kim B., Summers, Lawrence H., Holt, Charles C., Hall, Robert E., Baily, Martin Neil, & Clark, Kim B. 1979. Labor Market Dynamics and Unemployment: A Reconsideration. *Brookings Papers on Economic Activity*, **1979**(1).

Coleman, J. S. 1966. Equality of Educational Opportunity. *Washington DC, US dept. of Health, Education and Welfare, Office for Education.*

Corak, Miles. 2006. Do Poor Children Become Poor Adults? Lessons from a Cross-Country Comparison of Generational Earnings Mobility. *In: Dynamics of Inequality and Poverty (Research on Economic Inequality, Volume 13).* Emerald Group Publishing Limited.

Corak, Miles, & Piraino, Patrizio. 2011. The Intergenerational Transmission of Employers. *Journal of Labor Economics*, **29**(1).

Couch, Kenneth A., & Dunn, Thomas A. 1997. Intergenerational Correlation of Labor Market Status: A Comparison of the United States and Germany. *Journal of Human Resources*, **32**(1).

Couch, Kenneth A., & Placzek, Dana W. 2010. Earnings Losses of Displaced Workers Revisited. *American Economic Review*, **101**(1).

Cunha, Flavio, & Heckman, James J. 2007. The Technology of Skill Formation. *American Economic Review.*

Cunha, Flavio, Heckman, James J., & Schennach, Susanne M. 2010. Estimating the Technology of Cognitive and Noncognitive Skill Formation. *Econometrica*, **78**(3).

Erosa, Andrès, Fuster, Luisa, & Kambourov, Guergoui. 2012. Towards a micro-founded theory of aggregate labor supply. *Working Paper*.

FERNÁNDEZ-BLANCO, JAVIÉR, & PREUGSCHAT, EDGAR. 2014. Unemployment Duration, Exit Rates and Re-employment Wage. *Working Paper*.

FLYNN, JAMES R. 2009. *What is Intelligence? Beyond the Flynn Effect*. Cambridge University Press.

FONTAINE, FRANÇOIS. 2008. Why Are Similar Workers Paid Differently? The Role of Social Networks. *Journal of Economic Dynamics and Control*, **32**(12).

GALENIANOS, MANOLI. 2014. Hiring through referrals. *Journal of Economic Theory*, **152**.

GIBBONS, ROBERT, & KATZ, LAWRENCE F. 1991. Layoffs and Lemons. *Journal of Labor Economics*, **9**(4).

GONZALEZ, FRANCISCO M., & SHI, SHOUYONG. 2010. An Equilibrium Theory of Learning, Search and Wages. *Econometrica*, **78**(2).

GRANOVETTER, MARK S. 1973. The Strength of Weak Ties. *The American Journal of Sociology*, **78**(6).

GRIFFEN, ANDREW. 2011. Evaluating the Effects of Child Care Policies on Children's Cognitive Development and Maternal Labor Supply. *Working Paper*.

GURYAN, JONATHAN, HURST, ERIK, & KEARNEY, MELISSA. 2008. Parental Education and Parental Time with Children. *Journal of Economic Perspectives*, **22**(3).

GUVENEN, FATIH, & KURUSCU, BURHANETTIN. 2009. A Quantitative Analysis of the Evolution of the U.S. Wage Distribution: 1970-2000. *NBER Macroeconomics Annual*, **24**.

HALL, ROBERT E., & MILGROM, PAUL R. 2007. The Limited Influence of Unemployment on the Wage Bargain. *American Economic Review*, **98**(4).

HANUSHEK, ERIC A., & WOESSMANN, LUDGER. 2008. The Role of Cognitive Skills in Economic Development. *Journal of Economic Literature*.

HECKMAN, JAMES J., & BORJAS, GEORGE J. 1980. Does unemployment cause future unemployment? Definitions, questions and answers from a continuous time model of heterogeneity and state dependence. *Economica*, **47**(187).

Heckman, James J., Stixrud, Jora, & Urzua, Sergio. 2006. The effects of cognitive and noncognitive abilities on labor market outcomes and social behavior. *Journal of Labor Economics*, **24**(2).

Heckman, James J., Moon, Seong Hyeok, Pinto, Rodrigo, Savelyev, Peter A., & Yavitz, Adam. 2010a. The rate of return to the HighScope Perry Preschool Program. *Journal of Public Economics*, **94**(1-2).

Heckman, James J., Malofeeva, Lena, Pinto, Rodrigo, & Savelyev, Peter A. 2010b. Understanding the Mechanisms through which an Influential Early Childhood Program Boosted Adult Outcomes. *Unpublished Manuscript.*

Helmers, Christian, & Patnam, Manasa. 2011. The formation and evolution of childhood skill acquisition: Evidence from India. *Journal of Development Economics.*

Herault, Nicolas, & Kalb, Guyonne R.J. 2009. Intergenerational Correlation of Labour Market Outcomes. *Working Paper.*

Holzer, Harry. 1988. Search method use by unemployed youth. *Journal of Labor Economics*, **6**(1).

Hornstein, Andreas. 2012. Accounting for Unemployment: the Long and Short of It. *Working Paper.*

Huggett, Mark, Ventura, Gustavo, & Yaron, Amir. 2011. Sources of Lifetime Inequality. *American Economic Review*, **101**(7).

Ioannides, Yannis, & Loury, Linda Datcher. 2004. Job information networks, neighborhood effects, and inequality. *Journal of Economic Literature*, **12**.

Jäntti, Markus, Bratsberg, Bernt, Røed, Knut, Raaum, Oddbjørn, Naylor, Robin, Österbacka, Eva, Björklund, Anders, & Eriksson, Tor. 2006. American exceptionalism in a new light: a comparison of intergenerational earnings mobility in the Nordic countries, the United Kingdom and the United States. *Discussion Paper 1938, Institute for the Study of Labor (IZA).*

Jovanovic, Boyan. 1979. Job Matching and the Theory of Turnover. *Journal of Political Economy*, **87**(5).

Kahn, Lisa B. 2010. The long-term labor market consequences of graduating from college in a bad economy. *Labour Economics*, **17**(2).

Kahn, Lisa B., & Lange, Fabian. 2013. Employer Learning, Productivity and the Earnings Distribution: Evidence from Performance Measures. *Review of Economic Studies, Forthcoming.*

Keane, Michael P. 2011. Labor Supply and Taxes: A Survey. *Journal of Economic Literature*, **49**(4).

Keane, Michael P., & Wolpin, Kenneth I. 1996. The Career Decisions of Young Men. *Journal of Political Economy*, **105**(3).

Khan, Adnan Q., & Lehrer, Steven F. 2013. The Impact of Social Networks on Labour Market Outcomes: New Evidence from Cape Breton. *Canadian Public Policy*, **39**(1).

Kramarz, Francis, & Skans, Oskar Nordström. 2014. When Strong Ties are Strong: Networks and Youth Labor Market Entry. *The Review of Economic Studies*, **81**(3).

Lange, Fabian. 2007. The Speed of Employer Learning. *Journal of Labor Economics*, **25**(1).

Lee, Sang Yoon, & Seshadri, Ananth. 2012. *What is Intelligence? Beyond the Flynn Effect.* Cambridge University Press.

Ljunqvist, Lars, & Sargent, Thomas J. 1998. The European Unemployment Dilemma. *Journal of Political Economy*, **106**(3).

Lockwood, Ben. 1991. Information Externalities in the Labour Market and the Duration of Unemployment. *The Review of Economic Studies*, **58**(4).

Magruder, Jeremy R. 2010. Intergenerational Networks, Unemployment and Persistent Inequality in South Africa. *American Economic Journal: Applied Economics*, **2**(1).

McGrattan, Ellen R., & Rogerson, Richard. 2004. Changes in Hours Worked, 1950-2000. *Federal Reserve Bank of Minneapolis Quarterly Review*, **28**(1).

Menzio, Guido, & Shi, Shouyoung. 2011. Efficient Search on the Job and the Business Cycle. *Journal of Political Economy*, **119**(3).

Menzio, Guido, Telyukova, Irina, & Visschers, Ludo. 2012. Directed Search over the Life Cycle. *Working Paper.*

Michaud, Amanda. 2014. An Information Theory of Worker Flows and Wage Dispersion. *Working Paper.*

MICHELACCI, CLAUDIO, PIJOAN-MAS, JOSEP, & RUFFO, HERNAN. 2011. Inequality in Unemployment Risk and in Wages. *Working Paper.*

MONTGOMERY, JAMES D. 1991. Social Network and Labor Market Outcomes: Toward an Economic Analysis. *The American Economic Review*, **81**(5).

MORTENSEN, DALE T., & NAGYPAL, EVA. 2007. More on Unemployment and Vacancy Fluctuations. *Review of Economic Dynamics*, **10**(3).

MORTENSEN, DALE T., & PISSARIDES, CHRISTOPHER A. 1994. Job Creation and Job Destruction in the Theory of Unemployment. *Review of Economic Studies*, **61**(3).

MORTENSEN, DALE T., & VISHWANATH, TARA. 1994. Personal Contacts and Earnings. *Labour Economics*, **1**.

MUNSHI, KAIVAN. 2003. Networks in the Modern Economy: Mexican Migrants in the U.S. Labor Market. *Quarterly Journal of Economics*, **118**(2).

O'REGAN, KATHERINE M., & QUIGLEY, JOHN M. 1993. Family Networks and Youth Access to Jobs. *Journal of Urban Economics*, **34**.

OREOPOULOS, PHILIP, VON WACHTER, TILL, & HEISZ, ANDREW. 2012. The Short- and Long-Term Career Effects of Graduating in a Recession. *American Economic Journal: Applied Economics*, **4**(1).

OSUNA, VICTORIA, & RIOS-RULL, JOSÈ VICTOR. 2003. Implementing the 35 hours workweek by Means of Overtime Taxation. *Review of Economic Dynamics*, **6**(1).

PELLIZZARI, MICHELE. 2010. Do friends and relatives really help in getting a good job? *Industrial and Labor Relations Review*, **63**(3).

PISTAFERRI, LUIGI. 1999. Informal networks in the Italian labor market. *Giornale degli Economisti e Annali di Economia*, **58**(3-4).

PISTAFERRI, LUIGI. 2003. Anticipated and Unanticipated Wage Changes, Wage Risk, and Intertemporal Labor Supply. *Journal of Labor Economics*, **21**(3).

RAMEY, GAREY, & RAMEY, VALERIE A. 2010. The Rug Rat Race - Comments and Discussion. *Brookings Papers on Economic Activity.*

Rees, Albert. 1966. Information Networks in Labor Markets. *The American Economic Review*, **56**(1/2).

Restuccia, Diego, & Urrutia, Carlos. 2004. Intergenerational Persistence of Earnings; The Role of Early and College Education. *The American Economic Review*, **94**(5).

Robinson, John P., & Godbey, Geoffrey. 1999. *Time for Life: The Surprising Ways Americans Use Their Time, 2nd Edition.* Pennsylvania State University Press.

Sandberg, John F., & Hofferth, Sandra L. 2001. Changes in children's time with parents: United States, 1981-1997. *Demography*, **38**(3).

Schmillen, Achim, & Möller, Joachim. 2012. Distribution and determinants of lifetime unemployment. *Labour Economics*, **19**(1).

Schmutte, Ian M. 2010. Job Referral Networks and the Determination of Earnings in Local Labor Markets. *Unpublished Paper.*

Schoelmann, Todd. 2012. Refugees and Early Childhood Human Capital. *Working Paper.*

Schwartz, Christine R., & Mare, Robert D. 2005. Trends in Educational Assortative Marriage from 1940 to 2003. *Demography*, **42**(4).

Shimer, Robert. 2005. The Cyclical Behavior of Unemployment and Vacancies. *American Economic Review*, **95**(1).

Shimer, Robert. 2008. The Probability of Finding a Job. *American Economic Review P&P*, **98**(2).

Shimer, Robert. 2012. Reassessing the Ins and Outs of Unemployment. *Review of Economic Dynamics*, **15**.

Solon, Gary. 1992. Intergenerational Income Mobility in the United States. *American Economic Review*, **82**(3).

Todd, Petra E., & Wolpin, Kenneth I. 2007. The production of cognitive achievement in children: Home, school, and racial test score gaps. *Journal of Human Capital.*

Topa, Giorgio. 2001. Social Interactions, Local Spillovers and Unemployment. *Review of Economic Studies*, **68**(2).

Topa, Giorgio. 2011. Labor Markets and Referrals. *In:* Elsevier, Amsterdam. (ed), *The Handbook of Social Economics*. North Holland.

von Wachter, Till, & Bender, Stefan. 2006. In the Right Place at the Wrong Time: The Role of Firms and Luck in Young Workers' Careers. *American Economic Review*.

von Wachter, Till, Song, Jae, & Manchester, Joyce. 2009. Long-Term Earnings Losses Due to Mass Layoffs During the 1982 Recession: An Analysis Using U.S. Administrative Data from 1974 to 2004. *Mimeo*.

Wiczer, David. 2014. Long Term Unemployment: Attached and Mismatched? *Working Paper*.